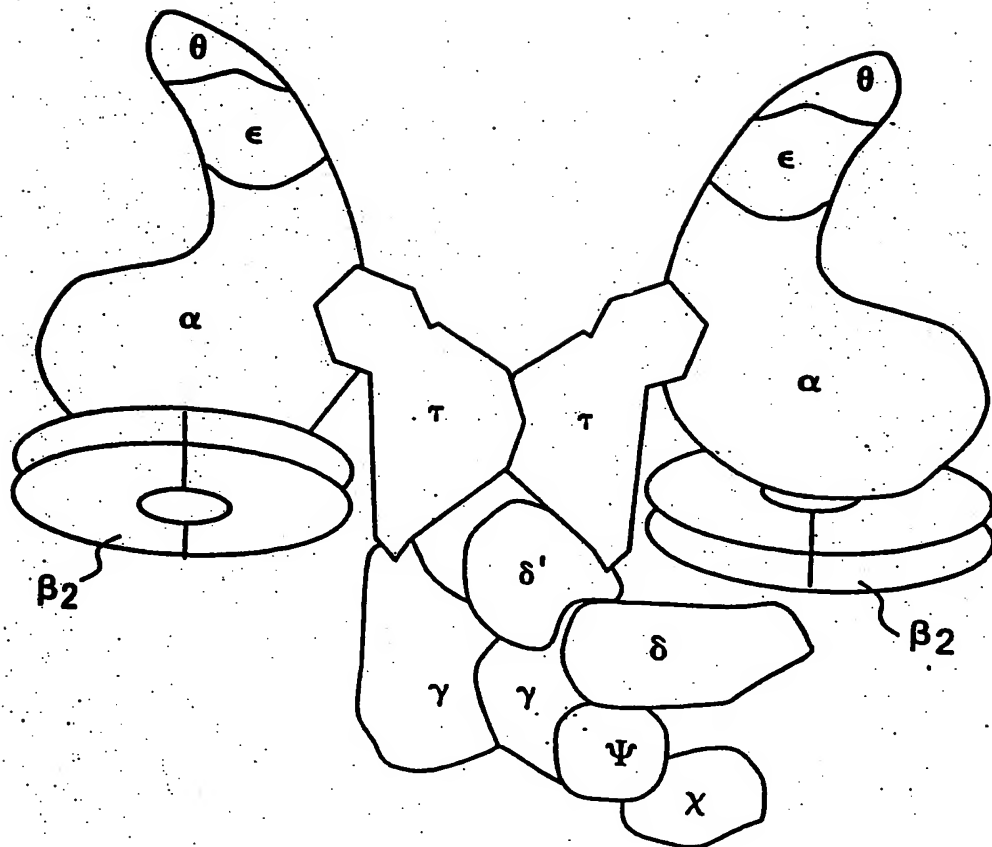


FIG.1



# ATP binding

E. coli  
MSYQVLARKWRPQTFADVVGQEHVLTALANGLSLGRIHHAYLFSGTRGVGKTSIARLLAK  
B. subtilis  
MSYQALYRVFRPQRFEDVVGQEHITKTLQNALLOKKFSHAYLFSGPRGTGKTSAAKIFAK  
\*\*\* \*

E. coli  
GLNCETGITATPCGVCDNCREIEQGRFVDLIEIDAASRTKVEDTRDLLDNVQYAPARGRF  
B. subtilis  
AVNCEHAPVDEPCNECAACKGITNGSISDVIEIDAASNNGVDEIRDIRDKVKFAPSAVTY  
\*\*\* \*

E. coli  
KVYLIDEVHMLSRHSFNALLKTLLEPPPEHVKFLATTDPOKL PVTILSRCLQFHLKALDV  
B. subtilis  
KVYIIDEVHMLSIGAFNALLKTLLEPPPEHCIFILATTEPHKIPLTIIISRCQRFDFKRITS  
\*\*\* \*

FIG. 2

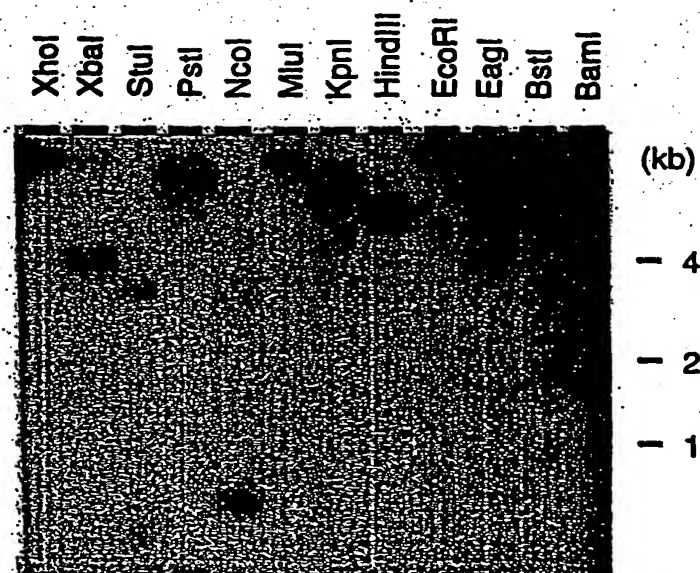


FIG.3

TCCGGGGGTG	GGGTTCCAG	GTAGACCCG	GCCCTCCCG	TGAGCCCCTT	TACCCAGGCC	60
GCCACCTCCT	CCAGGGGGC	CAAGCGTGC	AAGGAGAGGA	ACGTCCGCAC	<u>CAGGCCCTAT</u>	120
						S.D.
ACTAGCCTT	GTG AGC GCC CTC TAC CGC CGC TTC CGC CCC CTC ACC TTC CAG GAG GTG GTG					180
	met ser ala leu tyr arg phe arg pro leu thr phe gln glu val val					(17)
						CAC
GGG CAG GAG CAC GTG AAG GAG CCC CTC CTC AAG GCC ATC CGG GAG GGG AGG CTC GCC CAG						240
gly gln glu his val lys glu pro leu leu lys ala ile arg glu gly arg leu ala gln						(37)
GCS TAC CTS TTC TCC GGS AC						
GCC TAC CTC TTC TCC GGG CCC AGG GGC GTG GGC AAG ACC ACC AGG GCG AGG CTC CTC GCC						300
ala tyr leu phe ser gly pro arg gly val gly lys thr thr ala arg leu leu ala						(57)
ATG GCG GTG GGG TGC CAG GGG GAA GAC CCC CCT TGC GGG GTC TGC CCC CAC TGC CAG GCG						360
met ala val gly cys gln gly glu asp pro pro cys gly val cys pro his cys gln ala						(77)
GtG CAG AGG GGC GCC CAC CCG GAC GTG GTG GAC ATT GAC GCC AGC AAC AAC TCC GTG						420
val gln arg gly ala his pro asp val val asp ile asp ala ala ser asn ser val						(97)
GAG GAC GTG CGG GAG CTG AGG GAA AGG ATC CAC CTC GCC CCC CTC TCT GCC CCC AGG AAG						480
glu asp val arg glu leu arg glu arg ile his leu ala pro leu ser ala pro arg lys						(117)
						C
GTC TTC ATC CTG GAC GAG GCC CAC ATG CTC TCC AAA AGC GCC TTC AAC GCC CTC CTC AAG						540
val phe ile leu asp Glu ala his met leu ser lys ser ala phe asn ala leu leu lys						(137)

FIG.4A-1





GAG CGC CTC GCC CGC CGC TCC GAC GCC TTA AGC CTG GAG GTG GCC CTC CTG GAG GCG GGA	1140
glu arg leu ala arg arg ser asp ala leu ser leu glu val ala leu leu glu ala gly	(337)
AGG GCC CTG GCC GAG GCC CTA CCC CAG CCC ACG GGC GCT CCT TCC CCA GAG GTC GGC	1200
arg ala leu ala ala glu ala leu pro gln pro thr gly ala pro ser pro glu val gly	(357)
CCC AAG CCG GAA AGC CCC CCG ACC CCG GAA CCC CCA AGG CCC GAG GAG GCG CCC GAC CTG	1260
pro lys pro glu ser pro pro thr pro glu pro pro arg pro glu glu ala pro asp leu	(377)
CGG GAG CGG TGG CGG GCC TTC CTC GAG GCC CTC AGG CCC ACC CTA CGG GCC TTC GTG CGG	1320
arg glu arg trp arg ala phe leu glu ala leu arg pro thr leu arg ala phe val arg	(397)
GAG GCC CGC CGG GAG GTC CGG GAA GGC CAG CTC TGC CTC GCT TTC CCC GAG GAC AAG GCC	1380
glu ala arg pro glu val arg glu gly gln leu cys leu ala phe pro glu asp lys ala	(417)
TTC CAC TAC CGC AAG GCC TCG GAA CAG AAG GTG AGG CTC CTC CCC CTG GCC CAG GCC CAT	1440
phe his tyr arg lys ala ser glu gln lys val arg leu leu pro leu ala gln ala his	(437)
frameshift site	
TTC GGG GTG GAG GAG GTC CTC GTC CTG GAG GGA GAA AAA AAA AGC CTG AGC CCA AGG	1500
phe gly val glu glu val val leu val leu glu gly glu lys lys ser leu ser pro arg	(457)

FIG.4B-1

CCC CGC CCG GCC CCA CCT CCT GAA GCG CCC GCA CCC CCG GGC CCT CCC GAG GAG GAG GTA	1560
pro arg pro ala pro pro pro glu ala pro ala pro pro gly pro pro glu glu val	(477)
GAG GCG GAG GAA GCG GCG GAG GAG GCC CCG GAG GAG GCC TTG AGG CCG GTG GTC CGC CTC	1620
glu ala glu glu ala ala glu ala pro glu glu ala leu arg arg val val arg leu	(497)
CTG GGG GGG CCG GTG CTC TGG GTG CCG CCG CCC AGG ACC CCG GAG CCG CCG GAG GAG GAA	1680
leu gly gly arg val leu trp val arg arg pro arg thr arg glu ala pro glu glu	(517)
CCC CTG AGC CAA GAC GAG ATA GGG GGT ACT GGT ATA TAA	1740
pro leu ser gln asp glu ile gly thr gly ile *	(529)
CGACCTCGGA CAAGAGACCG TGGACAACAT CCTCAAGCGC CTCCGCCGTA TTGAGGCCCA	1820
GGTGCGGGGG CTCCAGAAGA TGGTGGCCGA GGGCCGCCCC TGCACGAGG TCCTCACCCA	1880
GATGACCGCC ACCAAGAAGG CCATGGAGGC GCGGCCACC CTGATCCTCC ACGAGTTCCT	1940
GAACGTCTGC GCCGCCGAGG TCTCCGAGGG CAAGGTGAAC CCAAGAAGC CCGAGGAGAT	2000
CGCCACCATG CTGAAGAACT TCATCTA	2027

FIG.4B-2

GGG	CAG	GAG	CAC	GTG	AGC	GCC	CTC	TAC	CGC	CGC	TTC	CGC	CCC	CTC	ACC	TTC	CAG	GAG	GTG	GTG	51
GCC	TAC	CTC	TTC	TCC	GTG	AAG	CCC	GAG	CTC	CTC	AAG	GCC	ATC	CGG	GAG	GGG	AGG	CTC	GCC	CAG	111
ATG	GCG	GTG	GGG	TGC	CAG	GGG	GAA	GAC	CCC	GTG	CCT	TGC	GGG	GTG	ACC	CGG	AGG	CTC	CTC	GCC	171
GtG	CAG	AGG	GGC	GCC	CAC	CCG	GAC	GTG	GAC	ATT	GAC	GCC	GAC	GCC	GCC	AGC	AAC	AAC	TCC	GTC	231
GAG	GAC	GTG	CGG	GAG	CTG	AGG	GAA	AGG	ATC	CAC	CAC	CTC	GCC	CCC	CTC	TCT	GCC	CCC	AGG	AAG	291
GTC	TTC	ATC	CTG	GAC	GCC	CAC	ATG	ATG	CTC	TCC	AAA	AGC	AGC	GCC	TTC	AAC	GCC	CTC	CTC	AAG	351
ACC	CTG	GAG	GAG	CCC	CCC	CCC	GTC	GTC	CTC	TTC	GTG	TTC	TTC	GCC	ACC	ACC	GAG	CCC	GAG	AGG	411
ATG	CCC	CCC	ACC	ATC	ATC	CTC	TCC	CGC	ACC	CAG	CAC	TTC	CGC	TTC	CGC	CGC	CTC	ACG	GAG	AGG	471
GAG	ATC	GCC	TTT	AAG	CTC	CTC	CGG	CGC	ATC	CTG	GAG	GCC	GTG	GGG	CGG	GAG	GCG	GAG	GAG	GAG	531
GCC	CTC	CTC	CTC	CTC	CTC	GCC	CGC	CTG	GCG	GAC	GGG	GCC	CTT	AGG	GAC	GCG	GAA	AGC	CTC	CTG	591
GAG	GCG	TTC	CTC	CTC	CTC	CTG	GAA	GGC	CCC	CTC	ACC	CGG	AAG	GAG	GTG	GAG	CGC	GCC	CTA	GCC	651
TCC	CCC	CCA	GGG	ACC	GGG	GTC	GCC	ATC	GCC	ATC	GCC	GCC	TCC	CTC	GCG	AGG	GGG	AAA	ACG	GCG	711
GAG	GCC	CTG	GGC	CTC	CTC	GCC	CGG	CGC	CTC	TAC	GGG	GAA	GGG	TAC	GCC	CCG	AGG	AGC	CTG	GTC	771
TCC	GGC	CTT	TTG	GAG	GTC	GGG	GTG	GGC	GAA	GGC	CTC	TAC	GCC	GCC	TTC	GGC	CTC	GCG	GGA	ACC	831
CCC	CTT	CCC	GCC	CCG	CCG	CCC	CAG	GCC	CTG	ATC	GCC	ATC	ATG	ACC	GCC	CTG	GAC	GAG	GCC	ATG	891
GAG	GCG	CTC	GCC	CGC	GCC	CGC	GAC	GCC	CTG	GAC	AGC	CTG	GAG	GTG	GCC	CTC	CTG	GAG	GCC	GGA	951
AGG	GCC	CTG	GCC	GCC	GAG	CCC	CTA	CCC	CGG	ACC	CCC	AGC	GGC	GCT	CCC	TCC	CCA	GAG	GTC	GCC	1011
CCC	AAG	CCG	GAA	AGC	AGC	CCC	ACC	ACC	CGG	GAA	CCC	CCA	AGG	AGG	CCC	GAG	GCG	CCC	GAC	CTG	1071
CGG	GAG	CGG	TGG	CGG	GCC	TTC	CTC	CTC	GAG	GCC	CTC	AGG	CCC	ACC	CTA	CGG	GCC	TTC	GTG	CGG	1131
GAG	GCC	CGC	CGC	GAG	GTC	CGG	GAA	GGC	CAG	CGC	CTC	TGC	CTC	GCT	TTC	CCC	GAG	GAC	AAG	GCC	1191
TTT	CAC	TAC	CGC	AAG	GCC	TCC	GAA	CAG	AAG	GTG	AGG	CTC	CTC	CTC	CCC	CTG	GCC	CAG	GCC	CAT	1251
TTT	GGG	GTG	GAG	GAG	GTC	GTG	CTC	CTC	GTC	CTG	GAG	GGA	GAA	AAA	AGC	CTG	AGC	CTG	AGC	AGG	1311
CCC	GCG	CCG	GCC	CCA	CCT	CCT	GAA	GCG	CCC	GCA	CCC	CCG	GGC	CCT	CCC	GAG	GAG	GAG	GTA	GTA	1371
GAG	GCG	GAG	GAA	GCG	GCG	GCG	GAG	GAG	GCC	CCG	GAG	GAG	GCC	TTG	AGG	CGG	GTG	GTC	CGC	CTC	1431
CTG	GGG	GGG	CGG	GTG	CTC	TGG	GTG	CGG	CGG	CCC	AGG	ACC	CGG	GAG	GCG	GAG	CGG	CCG	GAG	GAG	1491
CCC	CTG	AGC	CAA	GAC	GAG	ATA	GGG	GGT	ACT	GGT	ATA	TAA	(1590)								1551

FIG.4C

Met	ser	ala	leu	tyr	arg	arg	phe	arg	pro	leu	thr	phe	gln	glu	val	gly	gln	glu	20	
his	val	lys	glu	pro	arg	leu	lys	ala	ile	arg	glu	gly	arg	leu	ala	gln	ala	tyr	leu	40
phe	ser	gly	pro	arg	gly	val	gly	lys	thr	thr	thr	ala	arg	leu	ala	met	ala	val		60
gly	cys	gln	gly	glu	asp	pro	pro	cys	gly	val	cys	pro	his	cys	gln	ala	val	gln	arg	80
gly	ala	his	pro	asp	val	val	asp	ile	asp	ala	ala	ser	asn	asn	ser	val	glu	asp	val	100
arg	glu	leu	arg	glu	arg	ile	his	leu	ala	pro	leu	ser	ala	pro	arg	lys	val	phe	ile	120
leu	asp	glu	ala	his	met	leu	ser	lys	ser	ala	phe	asn	ala	leu	leu	lys	thr	leu	glu	140
glu	pro	pro	pro	his	val	leu	phe	val	phe	ala	thr	thr	glu	pro	glu	arg	met	pro	pro	160
thr	ile	leu	ser	arg	thr	gln	his	phe	arg	phe	arg	arg	leu	thr	glu	glu	glu	ile	ala	180
phe	lys	leu	arg	arg	ile	leu	glu	ala	val	gly	arg	glu	ala	glu	glu	glu	ala	leu	leu	200
leu	leu	ala	arg	leu	ala	asp	gly	ala	leu	arg	asp	ala	glu	ser	leu	leu	glu	arg	phe	220
leu	leu	leu	glu	gly	pro	leu	thr	arg	lys	glu	val	glu	arg	ala	leu	gly	ser	pro	pro	240
gly	thr	gly	val	ala	glu	ile	ala	ala	ser	leu	ala	arg	gly	lys	thr	ala	glu	ala	leu	260
gly	leu	ala	arg	arg	leu	tyr	gly	glu	gly	tyr	ala	pro	arg	ser	leu	val	ser	gly	leu	280
leu	glu	val	phe	arg	glu	gly	leu	tyr	ala	ala	phe	gly	leu	ala	gly	thr	pro	leu	pro	300
ala	pro	pro	gln	ala	leu	ile	ala	ala	met	thr	ala	leu	asp	glu	ala	met	glu	arg	leu	320
ala	arg	ser	ser	asp	ala	leu	ser	leu	glu	val	ala	leu	leu	glu	ala	gly	arg	ala	leu	340
ala	ala	glu	ala	leu	pro	gln	pro	thr	gly	ala	pro	ser	pro	glu	val	gly	pro	lys	pro	360
glu	ser	pro	pro	thr	pro	glu	pro	pro	arg	pro	glu	glu	ala	pro	asp	leu	arg	glu	arg	380
trp	arg	ala	phe	leu	glu	ala	leu	ala	leu	thr	leu	arg	ala	phe	val	arg	glu	ala	arg	400
pro	glu	val	arg	glu	gly	gln	lys	val	leu	ala	phe	pro	glu	asp	lys	ala	phe	his	tyr	420
arg	lys	ala	ser	glu	gln	lys	val	arg	leu	leu	pro	leu	ala	gln	ala	his	phe	gly	val	440
glu	glu	val	val	leu	val	leu	glu	gly	glu	lys	lys	ser	leu	ser	pro	arg	pro	arg	pro	460
ala	pro	pro	pro	glu	ala	pro	ala	pro	pro	gly	pro	pro	glu	glu	glu	val	glu	ala	glu	480
glu	ala	ala	glu	glu	ala	pro	glu	glu	ala	leu	arg	arg	val	val	arg	leu	leu	gly	gly	500
arg	val	leu	trp	val	arg	arg	pro	arg	thr	arg	glu	ala	pro	glu	glu	pro	leu	ser		520
gln	asp	glu	ile	gly	gly	thr	gly	ile												529

FIG.4D

Met	ser	ala	leu	tyr	arg	arg	phe	thr	leu	thr	phe	gln	gln	val	gly	gln	glu	20			
his	val	lys	glu	pro	leu	lys	ala	ile	arg	glu	gly	arg	leu	ala	gln	ala	tyr	leu	40		
phe	ser	gly	pro	arg	gly	val	gly	lys	thr	thr	ala	arg	leu	ala	met	ala	val	60			
gly	cys	gln	gly	glu	asp	pro	pro	cys	gly	val	cys	pro	his	cys	gln	ala	val	gln	arg	80	
gly	ala	his	pro	asp	val	val	asp	ile	asp	ala	ala	ser	asn	ser	val	gln	asp	val	100		
arg	glu	leu	arg	glu	arg	ile	his	leu	ala	pro	leu	ser	ala	pro	arg	lys	val	phe	ile	120	
leu	asp	glu	ala	his	met	leu	ser	lys	ser	ala	phe	asn	ala	leu	leu	lys	thr	leu	glu	140	
glu	pro	pro	pro	his	val	val	phe	val	phe	ala	thr	thr	glu	pro	glu	arg	met	pro	pro	160	
thr	ile	leu	ser	arg	thr	gln	his	phe	arg	phe	arg	leu	arg	thr	glu	glu	glu	ile	ala	180	
phe	lys	leu	arg	arg	ile	leu	glu	ala	val	gly	arg	glu	ala	glu	glu	ala	leu	leu	200		
leu	leu	ala	arg	leu	ala	asp	gly	ala	leu	arg	asp	ala	glu	ser	leu	leu	glu	arg	phe	220	
leu	leu	leu	glu	gly	pro	leu	thr	arg	lys	glu	val	glu	arg	ala	leu	gly	ser	pro	pro	240	
gly	thr	gly	val	ala	arg	leu	ile	ala	ala	ser	leu	ala	arg	gly	lys	thr	ala	glu	ala	leu	260
gly	leu	ala	arg	arg	leu	tyr	gly	leu	gly	tyr	ala	pro	arg	ser	leu	val	ser	gly	leu	280	
leu	glu	val	phe	arg	glu	glu	gly	leu	tyr	ala	phe	gly	leu	ala	gly	thr	pro	leu	pro	300	
ala	pro	pro	gln	ala	ala	leu	ile	ala	ala	met	thr	ala	leu	asp	glu	ala	met	glu	arg	leu	320
ala	arg	arg	ser	asp	ala	leu	ser	leu	glu	val	ala	leu	leu	leu	glu	ala	gly	arg	ala	leu	340
ala	ala	glu	ala	leu	pro	gln	pro	thr	gly	ala	pro	ser	pro	pro	glu	val	gly	pro	lys	pro	360
glu	ser	pro	pro	thr	pro	glu	pro	pro	arg	pro	glu	ala	ala	pro	asp	leu	arg	glu	arg	380	
trp	arg	ala	phe	leu	glu	ala	leu	ala	arg	pro	thr	leu	arg	ala	phe	val	arg	glu	ala	arg	400
pro	glu	val	arg	glu	gly	gln	leu	leu	cys	leu	phe	pro	glu	asp	lys	ala	phe	his	tyr	420	
arg	lys	ala	ser	glu	gln	lys	val	arg	leu	leu	pro	leu	ala	gln	ala	his	phe	gly	val	440	
glu	glu	val	val	leu	val	leu	glu	lys	glu	lys	lys	pro	lys	pro	asp	pro	lys	ala	pro	460	
gly	pro	thr	ser																	464	

FIG. 4E

Met	ser	ala	leu	tyr	arg	arg	phe	leu	thr	phe	gln	glu	val	val	gly	gln	glu	20			
his	val	lys	glu	pro	leu	lys	ala	ile	arg	glu	gly	arg	leu	ala	gln	ala	tyr	leu	40		
phe	ser	gly	pro	arg	gly	val	gly	thr	thr	thr	ala	arg	leu	ala	met	ala	val	60			
gly	cys	gln	gly	glu	asp	pro	pro	cys	gly	val	cys	pro	his	cys	gln	ala	val	gln	arg	80	
gly	ala	his	pro	asp	val	asp	ile	asp	ala	ala	ser	asn	asn	ser	val	glu	asp	val	100		
arg	glu	leu	arg	glu	arg	ile	his	leu	ala	pro	leu	ser	ala	pro	arg	lys	val	phe	ile	120	
leu	asp	glu	ala	his	met	leu	ser	lys	ser	ala	phe	asn	ala	leu	leu	lys	thr	leu	glu	140	
glu	pro	pro	pro	his	val	leu	phe	val	phe	ala	thr	thr	glu	pro	glu	arg	met	pro	pro	160	
thr	ile	leu	ser	arg	thr	gln	his	phe	arg	phe	arg	leu	thr	thr	glu	glu	ile	ala	ala	180	
phe	lys	leu	arg	arg	ile	leu	glu	ala	val	gly	arg	glu	ala	glu	glu	ala	leu	leu	200		
leu	leu	ala	arg	leu	ala	asp	gly	ala	leu	arg	asp	ala	glu	ser	leu	leu	glu	arg	phe	220	
leu	leu	leu	glu	gly	pro	leu	thr	arg	lys	glu	val	glu	arg	ala	leu	gly	ser	pro	pro	240	
gly	thr	gly	val	ala	glu	ile	ala	ala	ser	leu	ala	arg	gly	lys	thr	ala	glu	ala	leu	260	
gly	leu	ala	arg	arg	leu	tyr	gly	gly	glu	tyr	ala	pro	arg	ser	leu	val	ser	gly	leu	280	
leu	glu	val	phe	arg	glu	gly	leu	tyr	ala	phe	gly	leu	ala	ala	gly	thr	pro	leu	pro	300	
ala	pro	pro	gln	ala	leu	ile	ala	ala	met	thr	ala	leu	asp	glu	ala	met	glu	arg	leu	320	
ala	arg	arg	ser	asp	ala	leu	ser	leu	glu	val	ala	leu	leu	glu	ala	gly	arg	ala	leu	340	
ala	ala	glu	ala	leu	pro	gln	pro	thr	gly	ala	pro	ser	pro	glu	val	gly	pro	lys	pro	360	
glu	ser	pro	pro	thr	pro	glu	pro	pro	arg	pro	glu	ala	pro	asp	leu	arg	glu	arg	380		
trp	arg	ala	phe	leu	glu	ala	leu	arg	pro	thr	leu	arg	ala	phe	val	arg	glu	ala	arg	400	
pro	glu	val	arg	glu	gly	gln	leu	leu	cys	leu	ala	phe	pro	glu	asp	lys	ala	phe	his	tyr	420
arg	lys	ala	ser	glu	gln	lys	val	arg	leu	leu	pro	leu	ala	gln	ala	his	phe	gly	val	440	
glu	glu	val	val	leu	val	leu	glu	gly	glu	lys	lys	lys	ala							454	

FIG.4F

		ATP site	
E.coli	MSYQVLARKWRPQTADVVGQEHVLTALANGLSLGRHHAYLFSGTRGVGKTSIARLLAK	60	
H.inf.	.....K.....II.....KDN.L.....F..	60	
B.sub.	....A.Y.VF...R.E.....ITKT.Q.A.LQKFS.....P.T....A.KIF..	60	
C.cres.	DA.T.....Y.R..E.LI...AMVRT...AF.T...A..FMLT.V.....TT.....R	113	
M.gen.	-MH..FYQ.Y..IN.KQTL...SIRKI.V.AINRDKLPNG.I...E..T...TF.KII...	59	
T.th.	--VSA.Y.RF..L..QE.....KEP.LKAIRE..LAQ.....P.....TT.....M	58	

	Zn <sup>++</sup> finger	
	* * *	
E.coli	GLNCET----GITATPCGVCDNCREIEQGRFVDLIEIDAASRTKVEDTRDLLDNVQYAPA	116
H.inf.	....VH----V.....E.E..KA....N.I.....E.....K.V	116
B.sub.	AV...H----APVDE..NE.AA.KG.TN.SIS.V.....NNG.DEI..IR.K.KF..S	116
C.cres.	A..Y..DTVK.PSVDLTTEGYH..S.IE..HM.VL.L.....DEM.E...G.R...V	173
M.gen.	AI..LN----WDQIDV.NS..V.KS.NTNSAI.IV.....KNGIN.I.E.VE..FNH.F	115
T.th.	AVG.QG-----EDP.....PH.QAVQR.AHP.VVD.....NNS...V.E.RERIHL..L	112

E.coli	RGRFKVYLIDEVHMLSRHSFNALLKTLEEPPEHVKFLLATDPQKLPVTILSRCLQFHLK	176
H.inf.	V.....IGA.....Y.....	176
B.sub.	AVTY...I.....IGA.....CI.I...E.H.I.L.I...QR.DF..	176
C.cres.	EA.Y...I.....TAA.....P.A..IF...EIR.V.....QR.D.R	233
M.gen.	TFKK...IL..A...TTQ.WGG.....S.PY.L.IFT..EFN.I.L.....QS.FF..	175
T.th.	SAPR...FIL..A....KSA.....P..L.VF...E.ERM.P.....TQH.RFR	172

FIG.5A



E.coli	ALDVEQIRHQLEHILNEEHIAHEPRALQOLLARAEGSLRDALSLTDQAIASGDQ--VST	234
H.inf.	...ET..SQH.A...TQ.N.PF.DP..VK..K..Q..I..S.....M..R.--.TN	234
B.sub.	RITSQA.VGRMNK.VDA.QLQV.EGS.EII.S..H.GM.....L....SFSGDI--LKV	234
C.cres.	RVEPDVLVKHFDR.SAK.GARI.MD..A.I.....V..G...L....VQTERGQT.TS	293
M.gen.	KITSDL.LER.ND.AKK.K.KI.KD..IKI.DLSQ.....G...L..LAI.LIVKKL.LL	235
T.th.	R.TE.E.AFK.RR..EAVGREA.EE..L....L.D.A...E..LERFLLLEGP---LTR	229
E.coli	QAVSAMLGTLDDDDQALSLVEAMVEANGERVMALINEAAARGIEWEALLVEMGLLHRAM	294
H.inf.	NV..N...L...NYSVDILY.LHQG...LL.RTLQRV.DAAGD.DK..G.CAEK...Q..L	294
B.sub.	EDALLIT.AVSQLYIGK.AKSLHDK.VSDALETL..LLQQ.KDPAK.IED.IFYFRDMLL	294
C.cres.	TV.RD...LA.RS.TIA.Y.HVMAGKTKDALEGFRALWGF.ADPVVMLDV.DHC.AS.V	353
M.gen.	MLKKHLISLIEMQN.L.L.KQFYQ.I	260
T.th.	KE.ERA...SPPGTGVAEIAASLARGKTAELG.ARRLYGE.YAPRS.VSGL.EVFREGLY	289

FIG.5B

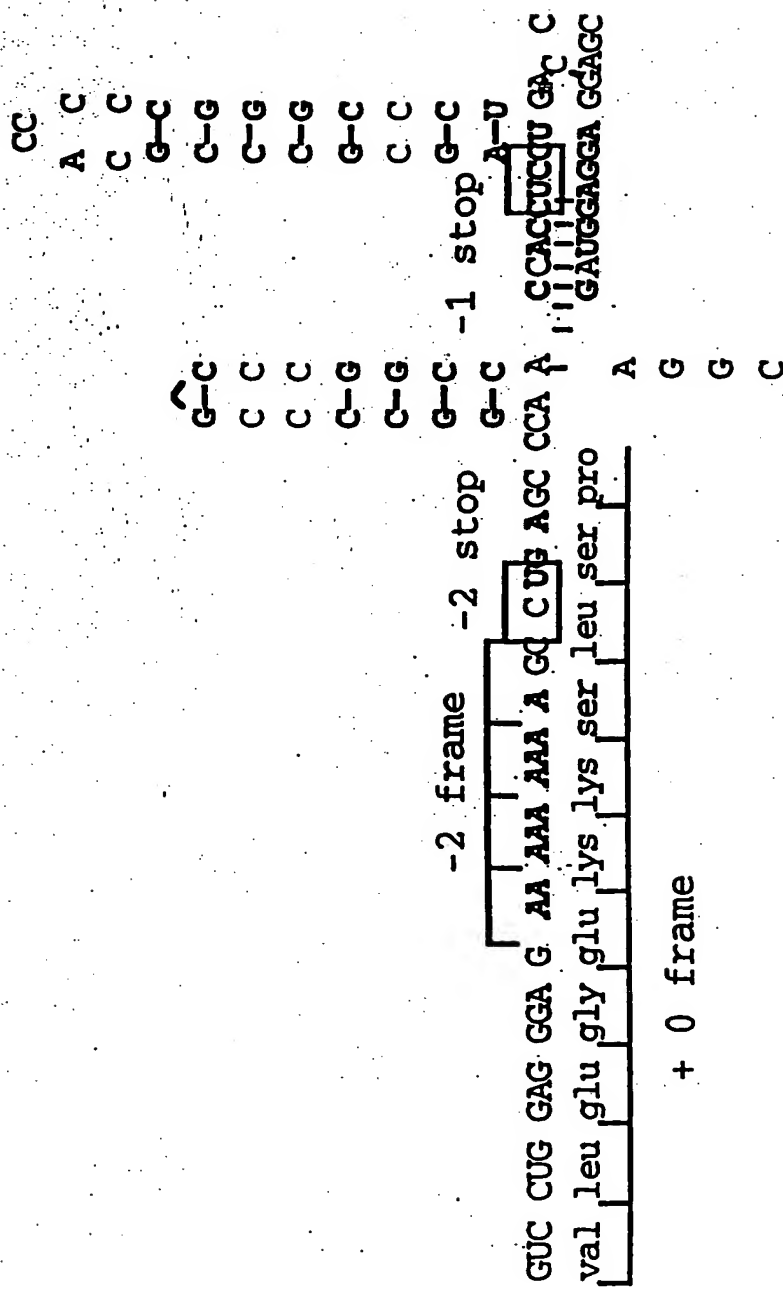


FIG.6

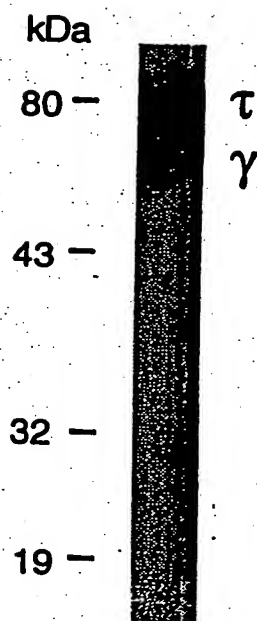
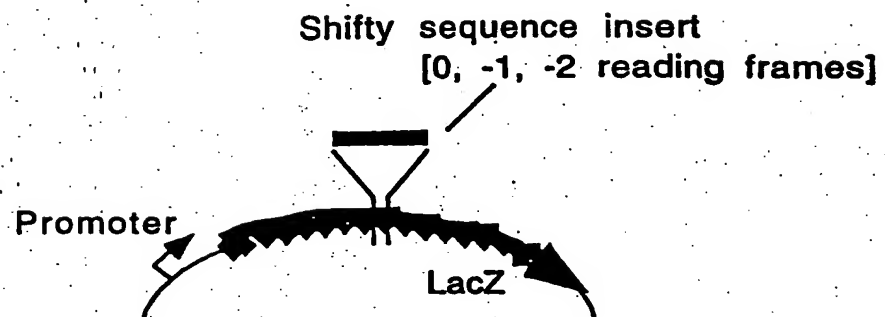


FIG.7

# FIG.8A



	Reading frame	Blue	White
Shifty sequence	0	+	
	- 1	+	
	- 2	+	
Mutant sequence	0	++	
	- 1		+
	- 2		+

# FIG.8B

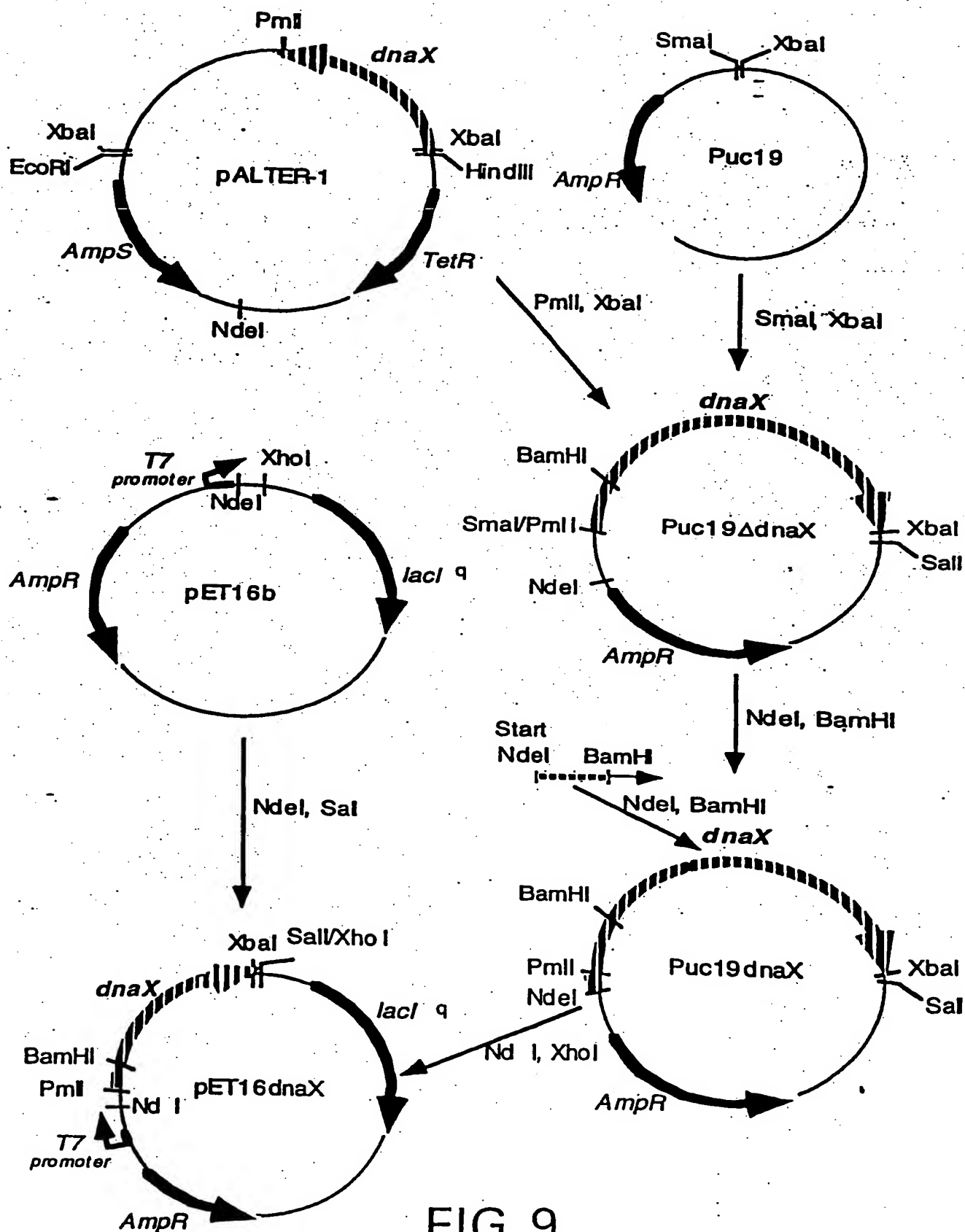


FIG.9

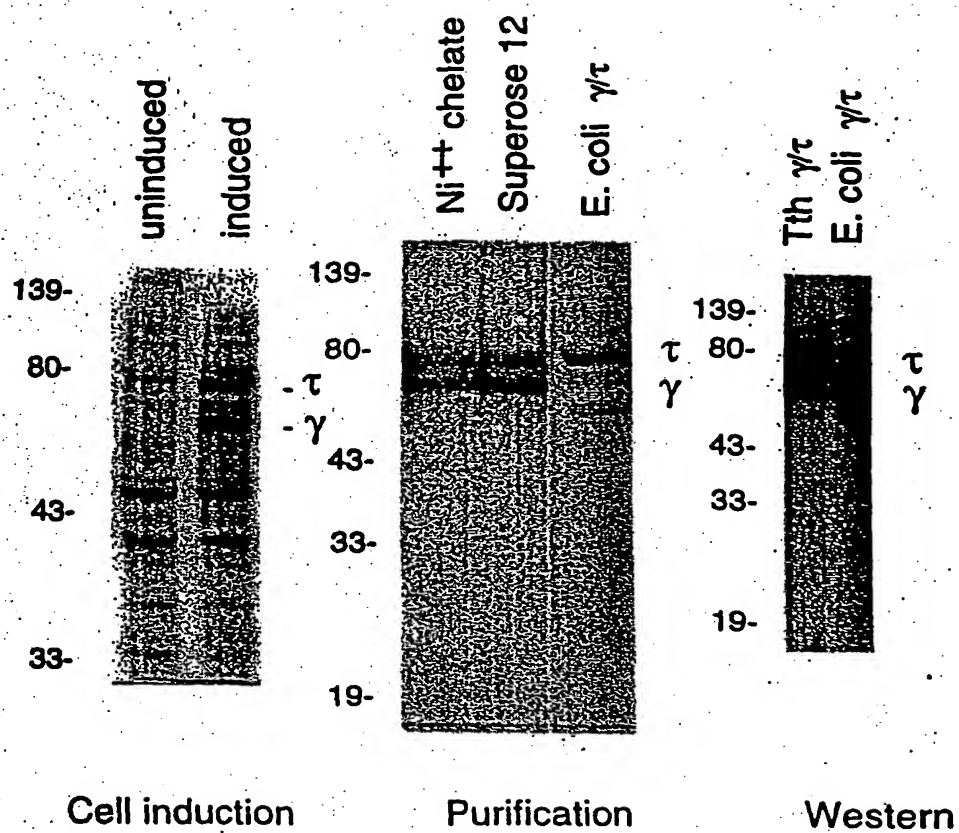


FIG.10A    FIG.10B    FIG.10C

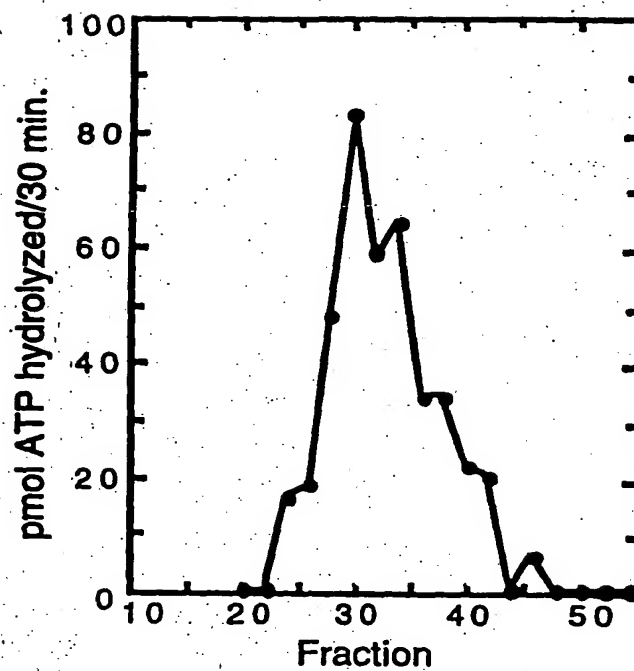


FIG. 11A

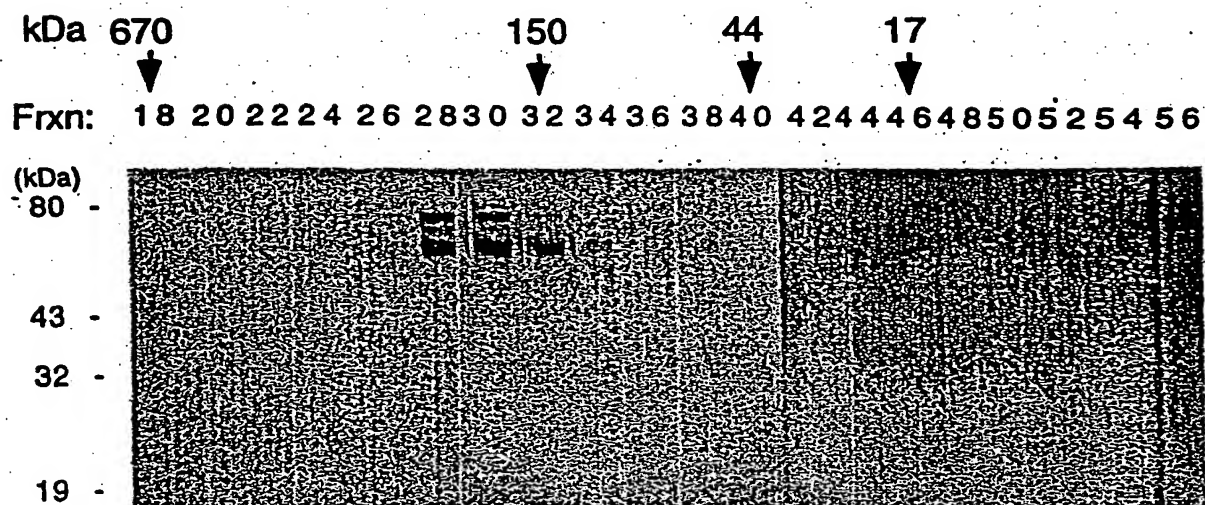


FIG. 11B

FIG.12A

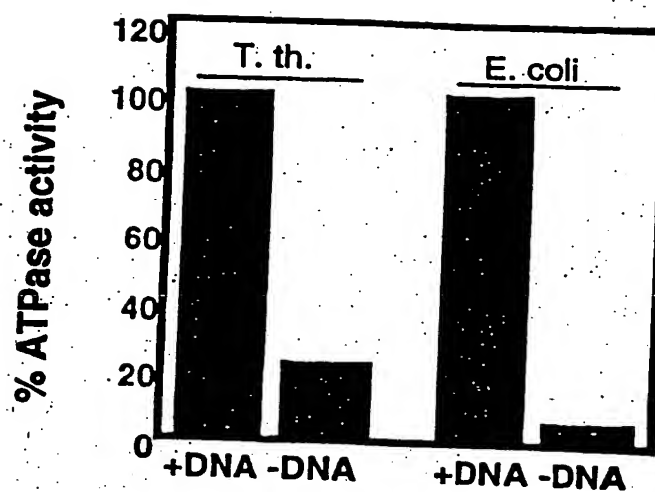


FIG.12B

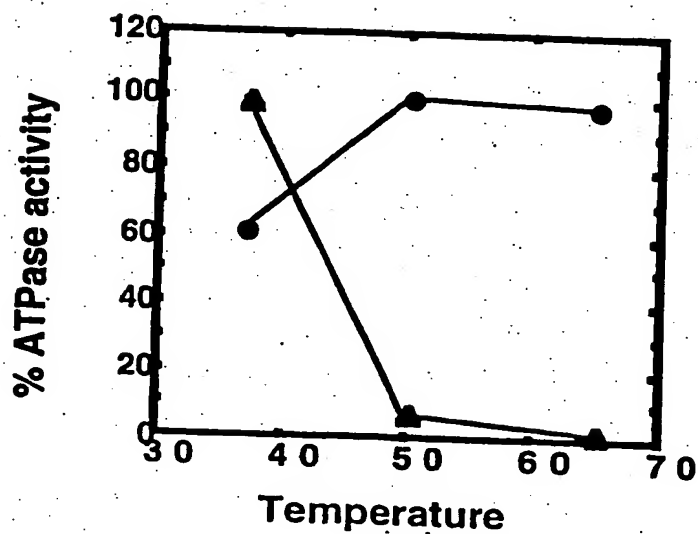
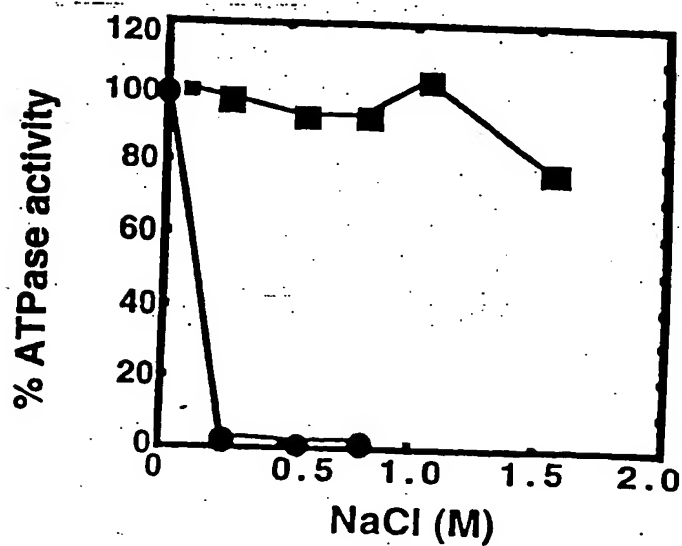
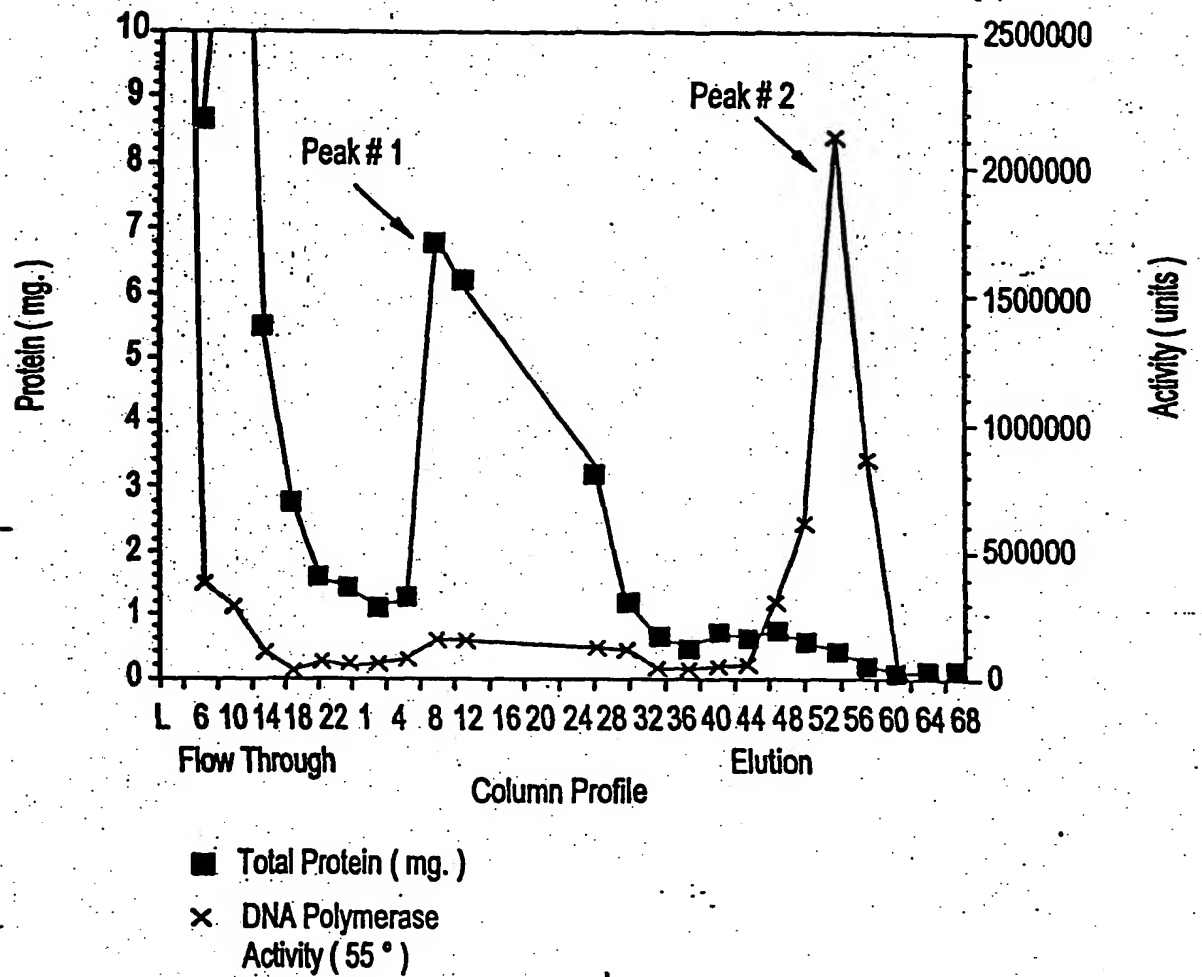


FIG.12C





# FIG.13A



# FIG.13B

ATP Agarose Step Column

FIG.13C

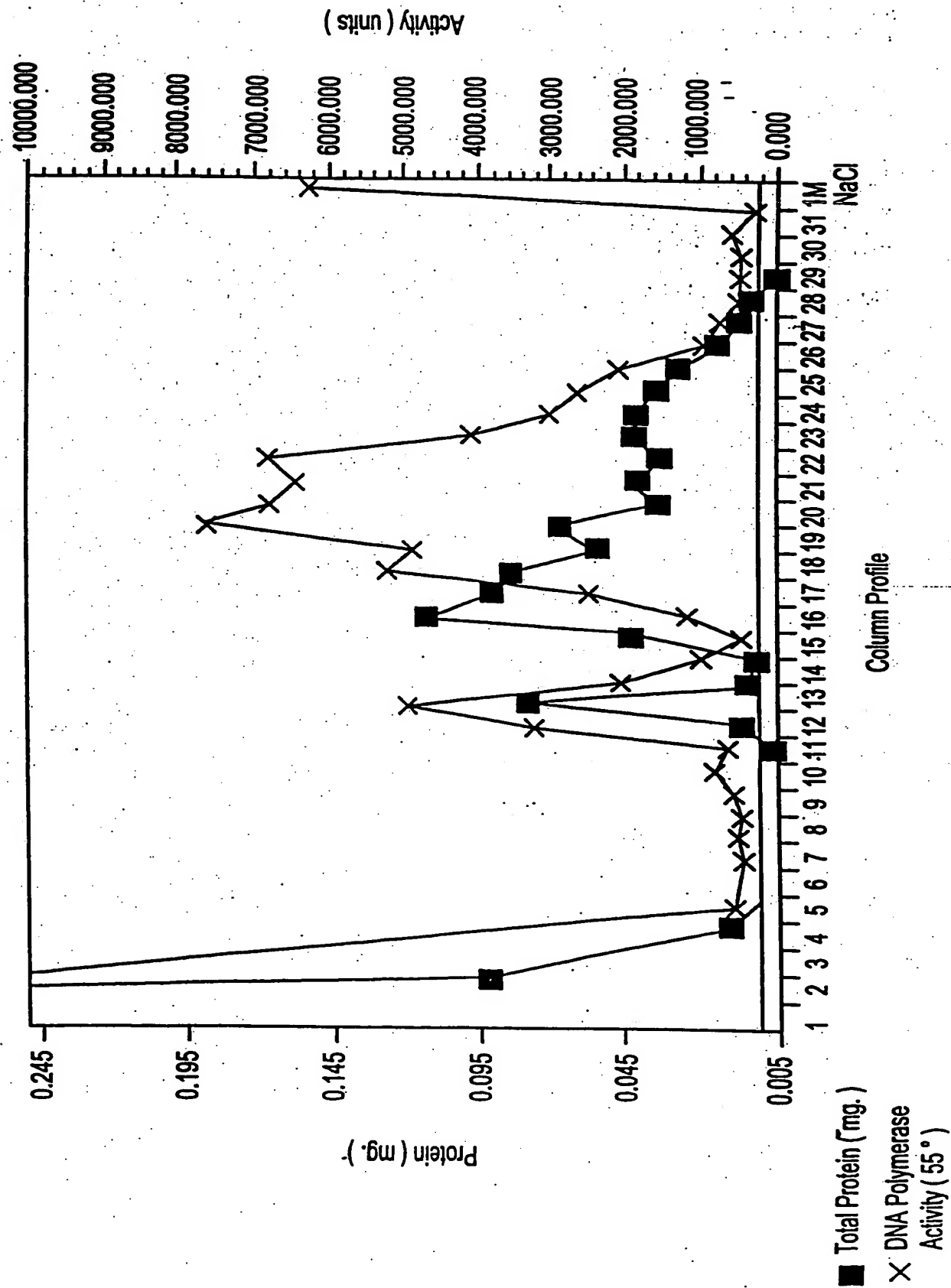
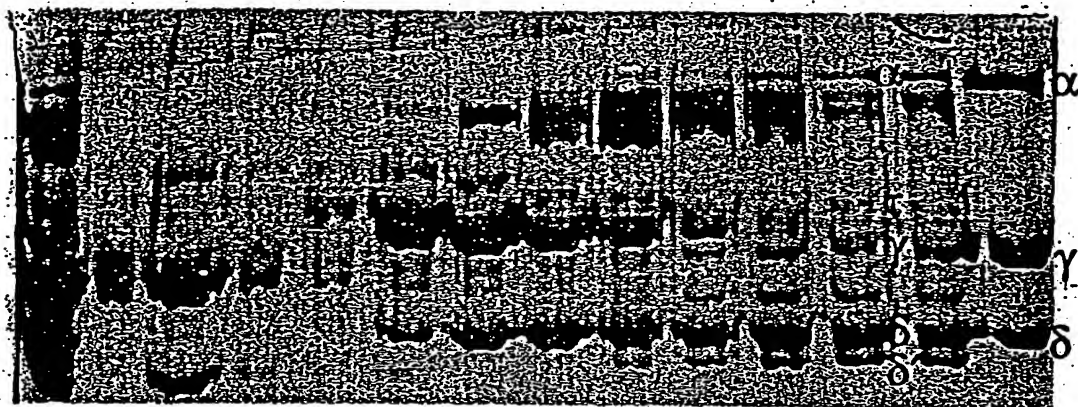


FIG.14A

load FT 9 10 11 12 13 14 15 16 17 18 19 E. coli  
 $\alpha$   $\gamma$   $\delta$



↑ ↑  
 T.th E. coli  
 subunits subunits

FIG.14B

load FT 9 10 11 12 13 14 15 16 17 18 19



# Alignment of TTH1 with alphas subunits of other organisms.

E.coli	DRYFLELIRTGRPDEESYLHAAVELAEARGLPV 197	(ID#72)
V.chol.	DHFYLELIRTGRADEESYLHFALDVAEQYDLPV 197	(ID#73)
H.inf.	DHFYLAISRTGRPNEERYIQALKLAERCDPLV 197	(ID#74)
R.prow.	DRFYFEIMRHDLPREEQFIENSYIQIASELSIPIV 195	(ID#75)
H.pyl.	DDFYLEIMRHGILDORFIDEQVIKMSLETGLKII 213	(ID#76)
S.sp.	DDYYLEIQDHGSVEDRLVNINLVKIAQELDIKIV 202	(ID#77)
M.tub.	DNYFLELMDHGLTIERRVRDGLLEIGRALNIPPL 220	(ID#78)
T.th.	FFIEIQNHGLSEQK	(ID#61)

FIG.15A

# Alignment of TTH2 with alphas subunits of other organisms.

E.coli	NKRRAKNGEPPLDIAAIPLDKKSFMDLQORSETTAVFQLESRGMKD 618	(ID#79)
V.chol.	NPRLLKAGKPPVRIEAIPLDDARSFRNLQDAKTTAVFQLESRGMK 618	(ID#80)
H.inf.	NVRMVREGKPRVDIAAIPLDDPESFELLKRSETTAVFQLESRGMKD 618	(ID#81)
R.prow.	CKKLLKEQGIKIDFDDMTFDDKKTYQMLCKGKGVGVFQFESIGMKD 624	(ID#82)
H.pyl.	LKIIKTQHKISVDFLSLDMDDPKVYKTIQSGDTVGFQIES-GMFQ 648	(ID#83)
S.sp.	QERKALQIRARTGSKKLPDDVKKTHKLLLEAGDLEGIFQLESQGMKQ 643	(ID#84)
M.tub.	IDNVRANRGIDLDESVPDLDKATYELLGRGDTLGVFQLDGGPMRD 646	(ID#85)
T.th.	RVELDYDALTLDD	(ID#60)

FIG.15B

ATGGGCCGGGAGCTCCGCTTCGCCCACCTCCACCAGCACA	
CCCAGTTCTCCCTCCTGGACGGGGCGGCGAAGCTTTCCGA	
CCTCCTCAAGTGGGTCAAGGAGACGACCCCCGAGGACCCC	120
GCCTTGCCATGACCGACCACGGCAACCTCTTCGGGGCCG	
TGGAGTTCTACAAGAAGGCCACCGAAATGGGCATCAAGCC	
CATCCTGGGCTACGAGGCCTACGTGGCGGCGGAAAGCCGC	240
TTTGACCGCAAGCGGGGAAAGGGCCTAGACGGGGGCTACT	
TTACCTCACCCTCCTCGCCAAGGACTTCACGGGGTACCA	
GAACCTGGTGCGCCTGGCGAGCCGGGCTTACCTGGAGGGG	360
TTTTACGAAAAGCCCCGGATTGACCGGGAGATCCTGCGCG	
AGCACGCCGAGGGCCTCATCGCCCTCTCGGGGTGCCTCGG	
GGCGGAGATCCCCCAGTTTCATCCTCCAGGACCGTCTGGAC	480
CTGGCCGAGGCCCGGCTCAACGAGTACCTCTCCATCTTCA	
AGGACCGCTTCTTCATCGAGATCCAGAACCACGGCCTCCC	
CGAGCAGAAAAAGGTCAACGAGGTCTCAAGGAGTTCGCC	600
CGAAAGTACGGCCTGGGGATGGTGGCCACCAACGACGGCC	
ATTACGTGAGGAAGGAGGACGCCCCGCGCCACGAGGTCCT	
CCTCGCCATCCAGTCCAAGAGCACCTTGACGACCCCGGG	720
CGCTGGCGCTTCCCCTGCGACGAGTTCTACGTGAAGACCC	
CCGAGGAGATGCGGGCCATGTTCCCCGAGGAGGAGTGGGG	
GGACGAGCCCTTTGACAACACCGTGGAGATCGCCCGCATG	840
TGCAACGTGGAGCTGCCCATCGGGGACAAGATGGTCTACC	
GAATCCCCCGCTTCCCCCTCCCCGAGGGGCGGACCGAGGC	
CCAGTACCTCATGGAGCTCACCTTCAAGGGGCTCCTCCGC	960
CGTACCCGGACCGGATCACCGAGGGCTTCTACCGGGAGG	
TCTTCCGCCTTTTGGGGAAGCTTCCCCCCCACGGGGACGG	
GGAGGCCTTGCCGAGGCCTTGGCCCAGGTGGAGCGGGAG	1080
GCTTGGGAGAGGCTCATGAAGAGCCTCCCCCTTTGGCCG	
GGGTCAAGGAGTGGACGGCGGAGGCCATTTTCCACCGGGC	
CCTTTACGAGCTTTCCGTGATAGAGCGCATGGGGTTTCCC	1200
GGCTACTTCTCATCGTCCAGGACTACATCAACTGGGCCC	
GGAGAAACGGCGTCTCCGTGGGGCCCCGGCAGGGGGAGCGC	
CGCCGGGAGCCTGGTGGCCTACGCCGTGGGGATCACCAAC	1320
ATTGACCCCTCCGCTTCGGCCTCCTCTTTGAGCGCTTCC	
TGAACCCGGAGAGGGTCTCCATGCCCGACATTGACACGGA	
CTTCTCCGACCGGGAGCGGGACCGGGTGATCCAGTACGTG	1440
CGGGAGCGCTACGGCGAGGACAAGGTGGCCAGATCGGCA	
CCCTGGGAAGCCTCGCCTCCAAGGCCGCCCTCAAGGACGT	
GGCCCCGGGTCTACGGCATCCCCACAAGAAGGCGGAGGAA	1560
TTGGCCAAGCTCATCCCGGTGCAGTTCGGAAGCCCAAGC	
CCCTGCAGGAGGCCATCCAGGTGGTGGCGGAGCTTAGGGC	
GGAGATGGAGAAGGACCCCAAGGTGCGGGAGGTCTTCGAG	1680
GTGGCCATGCGCCTGGAGGGCCTGAACCGCCACGCCTCCG	
TCCACGCCGCCGGGGTGGTGATCGCCGCCGAGCCCCCTCAC	
GGACCTCGTCCCCCTCATGCGCGACCAGGAAGGGCGGCCC	1800
GTACCCAGTACGACATGGGGGCGGTGGAGGCCTTGGGGC	
TTTTGAAGATGGACTTTTGGGCCTCCGCACCCTCACCTT	

FIG. 16A

CCTGGACGAGGTCAAGCGCATCGTCAAGGCGTCCCAGGGG	1920
GTGGAGCTGGACTACGATGCCCTCCCCCTGGACGACCCCA	
AGACCTTCGCCCTCCTCTCCCGGGGGGAGACCAAGGGGGT	
CTTCCAGCTGGAGTCGGGGGGGATGACCGCCACGCTCCGC	2040
GGCCTCAAGCCGCGGCGCTTTGAGGACCTGATCGCCATCC	
TCTCCCTCTACCGCCCCGGGCCCATGGAGCACATCCCCAC	
CTACATCCGCCGCCACCACGGGCTGGAGCCCCTGAGCTAC	2160
AGCGAGTTTCCCCACGCCGAGAAGTACCTAAAGCCCATCC	
TGGACGAGACCTACGGCATCCCCGTCTACCAGGAGCAGAT	
CATGCAGATCGCCTCGGCCGTGGCGGGGTACTCCCTGGGC	2280
GAGGCGGACCTCCTGCGGCGGTCCATGGGCAAGAAGAAGG	
TGGAGGAGATGAAGTCCACCGGGAGCGCTTCGTCCAGGG	
GGCCAAGGAAAGGGGCGTGCCCGAGGAGGAGGCCAACCGC	2400
CTCTTTGACATGCTGGAGGCCTTCGCCAACTACGGCTTCA	
ACAAATCCCACGCTGCCGCTACAGCCTCCTCTCCTACCA	
GACCGCCTACGTGAAGGCCCACTACCCCGTGGAGTTCATG	2520
GCCGCCCTCCTCTCCGTGGAGCGGCACGACTCCGACAAGG	
TGGCCGAGTACATCCCGACGCCCGGGCCATGGGCATAGA	
GGTCCTTCCCCCGGACGTCAACCGCTCCGGGTTTGACTTC	2640
CTGGTCCAGGGCCGGCAGATCCTTTTCGGCCTCTCCGCGG	
TGAAGAACGTGGGCGAGGCGGCGGCGGAGGCCATTCTCCG	
GGAGCGGGAGCGGGGCGGCCCTACCGGAGCCTCGGCGAC	2760
TTCTCAAGCGGCTGGACGAGAAGGTGCTCAACAAGCGGA	
CCCTGGAGTCCCTCATCAAGGCGGGCGCCCTGGACGGCTT	
CGGGGAAAGGGCGGGCTCCTCGCCTCCCTGGAAGGGCTC	2880
CTCAAGTGGGCGGCCGAGAACCGGGAGAAGGCCCGCTCGG	
GCATGATGGGCCTCTTCAGCGAAGTGAGGAGCCGCCTTT	
GGCCGAGGCCGCCCCCCCTGGACGAGATCACCCGGCTCCGC	3000
TACGAGAAGGAGGCCCTGGGGATCTACGTCTCCGGCCACC	
CCATCTTGCGGTACCCCGGGCTCCGGGAGACGGCCACCTG	
CACCCTGGAGGAGCTTCCCCACCTGGCCCGGGACCTGCCG	3120
CCCCGGTCTAGGGTCCTCCTTGCCGGGATGGTGGAGGAGG	
TGGTGCGCAAGCCCACAAAGAGCGGCGGGATGATGGCCCG	
CTTCGTCTCTCCGACGAGACGGGGGCGCTTGAGGCGGTG	3240
GCATTGCGCCGGGCCTACGACCAGGTCTCCCCGAGGCTCA	
AGGAGGACACCCCCGTGCTCGTCCTCGCCGAGGTGGAGCG	
GGAGGAGGGGGGCGTGCGGGTGCTGGCCCAGGCCGTTTGG	3360
ACCTACGAGGAGCTGGAGCAGGTCCCCCGGGCCCTCGAGG	
TGGAGGTGGAGGCCTCCCTCCTGGACGACCGGGGGGTGGC	
CCACCTGAAAAGCCTCCTGGACGAGCACGCGGGGACCTC	3480
CCCCTGTACGTCCGGGTCCAGGGCGCCTTCGGCGAGGCC	
TCCTCGCCCTGAGGGAGGTGCGGGTGGGGGAGGAGGCTGT	
AGGCGGCCGCGTGTTCCGGGCCTACCTCCTGCCCCGACCG	3600
GGAGGTCTTCTCCAGGGCGGCCAGGCGGGGGAGGCCAG	
GAGGCGGTGCCCTTCTAGGGGGTGGGCCGTGAGACCTAGC	
GCCATCGTTCTCGCCGGGGGCAAGGAGGCCTGGGCCCGAC	3720
CCCTTTTGG	

FIG. 16B

MGRELRF AHLHQHTQFSLLDGAPKLSDLLKWVEETTPEDP	
ALAMTDHGNLFGAVEFYKKATEMGIKPILGYEAYVAAESR	
FDRKRKGKLDGGYFHLTLLAKDFTGYQNLVRLASRAYLEG	120
FYEKPRIDREILREHAEGLIALSGCLGAEIPQFILQDRLD	
LAEARLNEYLSIFKDRFFIEIQNHGLPEQKKVNEVLKEFA	
RKYGLGMVATNDGHYVRKEDARAHEVLLAIQSKSTLDDPG	240
ALALPCEEFYVKTPEEMRAMFPEEEVGGRSPLTTPWRSPL	
VQGAAGTRWSTRI PRFPLPEGRTEAQYLMELTFKGLLR	
RYPDRITEGFYREVFRLSGKLPPHGDGEALAEALAQVERE	360
AWERLMKSLPPLAGVKEWTAEAI FHRALYELSAIERMGFP	
GLLPHRPGHLHQLGPEKGVSVGPGRGGAAGSLVAYAVGITN	
IDPLRFGLLFRFLNPERVSMPTDIDTDFSDRERDRVIQYV	480
RERYGEDKVAQIGTLGSLASKAALKEVARVYGI PRKKAEE	
LAKLIPVQFGKPKPLQEAIQVPELRAEMEKDPKVREVLE	
VAMRLEGLNRHASVHAGRGGVFSEPLTDLVPLCATRKGGP	600
YTQYDMGAVEALGLLKMDFLGLRTLTLFDEVKRIVKASQG	
VELDYDALPLDDPKTFALLSRGETKGVFQLESGGMTATLR	
GLKPRRFEDLIAILSLYRPGPMEHIPTYIRRHGPLEVSY	720
SEFPHAEKYLKPILDETYGIPVYQEQIMQIASAVAGYSLG	
EADLLRRSMGKKKVEEMKSHRERFVQAKERGVPEEEANR	
LFDMLEAFANYGFNKSHAAAYSLLSYQTAYVKAHYPVFEM	840
AALLSVERHDSKVAEYIRDARAMGIEVLPPDVNRSGFDF	
LVQGRQILFGLSAVKNVGEAAAEAILRERERGGPYRSLGD	
FLKRLDEKVLNKRITLES LIKAGALDGFGERARLLASLEGL	960
LKWAAENREKARSGMMGLFSEVEEPPLAEAAPLDEITRLR	
YEKEALGIYVSGHPILRYPGLRETATCTLEELPHLARDLP	
PRSRVLLAGMVEEVVRKPTKSGGMMARFVLSDETGALEAV	1080
AFGRAYDQVSPRLKEDTPVLVLAEVEREEGGVRVLAQAVW	
TYQELEQVPRALEVEVEASLPDDRGV AHLKSLLEHAGTL	
PLYVRVQGAFG EALLALREVRVGEEALGALEAAGFPAYLL	1200
PNREVSPRLTGSGGPRGRALSTGLALKTYP IALPGNEAL	
ARPLL	

FIG. 16C

	Start1	Start2	3'-Exo I
T.th.	VERVVRTLLDGRFLLEEGVGLWEWRYPPFLEGEAVVLDLLETTGLAG-----LDEVIEVGLLRLEGG---RRLPF		
D.rad.		PWPQDVVVDLLETTGSPA-----SAAIVEIGAVRIVGGQIDETLKF	
Bac.sub.	HGIKMIYMEANLVDDGVPPIAYNAAHRLLEEETVVFDDVETTGLSAV-----YDTIIELAAVKVKGGE--IIDKF		
H.inf.		MINPNRQIVLDTFETTGMMNQLGAHYEGHCHIIIEIGAVELINRR-YTGNNX	
E.c.		MSTAITRQIVLDTFETTGMMNQIGAHSEGHKIIIEIGAVEVNNRR-LTGNNF	
H.pyl.	NLEYLKACGLNFIEITSENLTITLKNLKTPLKDEVFSFIDLETTGSCPI-----KHEILLEIGAVQVKGGE--IINRF		

	3'-Exo II
T.th.	QSLVR-PLPP---AARSWNLT---GIPREALEEAPSLEEVLEKAYPLRGDATLVTHNAAFDLGFL-RPALEGLG
D.rad.	ETLVR-PTRPDGSMLSIPWQAQRVHGISDEMVRRAPAKDVLPDFDFVDGSAAVVAHNVSFDGGFM-RAGAERLG
Bac.sub.	EAFAN-PHRP---LSATIIELT---GITDDMLQADPDVVDVIRDFREWIGDDILVAHNASDFDMGFL-NVAYKKLL
H.inf.	HIYIK-PDRP---XDPDAIKVH---GITDEMLADKPEFKEVAQDFLDYINGAELLIHNAAPFDVGFM-DYEFKRKLN
E.c.	HVYLK-DRLV---DPEAFGVH---GIAVDFLLDKPTFAEVAVEFMDYIRGAELVTHNAAFDIDGFM-DYEFSLLK
H.pyl.	ETLVKVKSV-----DYIAELT---GITYEDTLNAPSAHEALQELRLFLGNSVFVAHNANFDYNFLGRYFVEKLH

	3'-Exo IIIC
T.th.	-----YRLENPVVDSLRLARRGLPGLRRYGLDALSEVLELPRRT--CHRALEDVERTLAVVHEVYVYMLT-----SG
D.rad.	-----LSWAPERELCTMQLSRRAFPFRERTHNLTVLAERLGLLEFAPGGRHRSYGDVQVTAQAYLRLLLELLG-----ER
Bac.sub.	E-----VEKAKNPVIDTLELGRFLYPEFKNHRLNTLCKKFDIELTQ--HHRATYDTEATAYLLLKMLKDA-----EK
H.inf.	-LNVKTDDICLVDTLQMARQMPGKRN-NLDALCDRLGIDNSKRTLHGALLDAEIIADVYLMMTGGQTNLFDEEE
E.c.	RDIAKTNTFCVKVTDLSAVARKMFPGKRN-SLDALCARYEIDNSKRTLHGALLDAQILAEVYLA MTGGQTSMAFAME
H.pyl.	-----CPLLNLKCLTDLKRAILSMRY-SLSFLKELLGFGIEV--SHRAYADALASYKLFELICLLNLP--SYIKT

FIG.17



## FIG.18A

ATGGTGGAGCGGGTGGTGCGGACCCTTCTGGACGGGAGGT 40  
TCCTCCTGGAGGAGGGGGTGGGGCTTTGGGAGTGGCGCTA  
CCCCTTTCCCCCTGGAGGGGGAGGCGGTGGTGGTCCTGGAC 120  
CTGGAGACCACGGGGCTTGCCGGCCTGGACGAGGTGATTG  
AGGTGGGCCTCCTCCGCCTGGAGGGGGGAGGCGCCTCCC 200  
CTTCCAGAGCCTCGTCCGGCCCCCTCCCGCCCGCCGAAGCC  
CGTTCGTGGAACCTCACCGGCATCCCCGGGAGGCCCTGG 280  
AGGAGGCCCCCTCCCTGGAGGAGGTTCTGGAGAAGGCCTA  
CCCCCTCCGCGGCGACGCCACCTTGGTGATCCACAACGCC 360  
GCCTTTGACCTGGGCTTCCTCCGCCCGGCCTTGGAGGGCC  
TGGGCTACCGCCTGGAAAACCCCGTGGTGGACTCCCTGCG 440  
CTTGGCCAGACGGGGCTTACCAGGCCTTAGGCGCTACGGC  
CTGGACGCCCTCTCCGAGGTCTTGGAGCTTCCCCGAAGGA 520  
CCTGCCACCGGGGCCCTCGAGGACGTGGAGCGCACCCCTCGC  
CGTGGTGCACGAGGTATACTATATGCTTACGTCCGGCCGT 600  
CCCCGCACGCTTTGGGAACTCGGGAGGTAG

MVERVVRTLLDGRFLLEEGVGLWEWRYPFPLEGEAVVLD 40  
LETTGLAGLDEVIEVGLLRLEGGRRLPFQSLVRPLPPAEA  
RSWNLTGIPREALLEEAPSLEEVLEKAYPLRGDATALVIHNA 120  
AFDLGFLRPALEGLGYRLENPVVDSLRLARRGLPGLRRYG  
LDALSEVLELPRRTCHRALEDVERTLAVVHEVYYMLTSGR 200  
PRTLWELGRZ

## FIG.18B

# Alignment of dnaA genes.

P. mar.	MLEASWEK VQSSL--KQNLK--	-----PSYE TWIRTEFSG--FKN GELTLIAPNSFSSAW LKNVYSQTIQETAE-	65
Syn. sp.	MVSCENLWQQ ALAIL--ATQLTK--	-----PAFD TWIKASVLIS--LGD GVATIQVENGFLVNH LQKSYGPLIMEVLT-	67
B. sut.	MENILDLMNQ ALAQI--EKKLSK--	-----PSFE TWKSTKAHS--LQG DTLTITAPNEFARDW LESRYLHLIADTIY-	67
M. tub.	MTDDPGSGFTTVWNA VVSELNGDPKVDIGP	SSDANLSAPLTPQOR AWLNLVQPLT--IVE GFALLSVPSFVQNE IERHLRAPITDALS-	87
T. th.	MSHEAVWQH VLEHI--RRSITE--	-----VEFH TWFERIRPLG--IRD GVLELAVPTSFALDW IRRHYAGLIOEGPR-	66
E. coli	MSLSLWQQ CLARL--QDELPA--	-----TEFS MWIRPLOAE--LSD NTLALYAPNRFVLDW VRDKYLNININGLLT-	64
T. mar.	MKER ILQEI--KTRVNR--	-----KSWE LWFSSFDVKS--IEG NKVVFSGNLFIKEM LEKKYYSVLSKAVK-	61
H. pyl.	MDTNNNIEKE ILALVKQNPVSL--	-----IEYE NYFSQKYNPNASKS DIAFFYAPNQVLCIT ITAKYGALLKEILLSQ	72
P. mar.	EIFG---EPVTVHVK VKANAESSDERHYSSA P-----	---ITPPLEASPGSV DSSGSSLRLSK--- -KTLPLILNLRYVFNK	130
Syn. sp.	DLTG---QEITVKLI TDGLEPHS---LIGQ E-----	---SSLPMETTP--- -KNATALNGKYTFESR	115
B. sut.	ELTG---EELSIFV IPQNDVEDFMKPKQ VKKAVKEDTSDFPQN	-----EIDDSAAARGDNQHS WFSYTFERPHNTDSA TAGVTSLNRRYTFDT	119
M. tub.	RRLGH-QIQLGVRIA PPATDEADDTTVPPS ENPATTSPDTTND	-----EIDDSAAARGDNQHS WFSYTFERPHNTDSA TAGVTSLNRRYTFDT	176
T. th.	LLGAQ-APRFELRV PGVVQEDIFQPPPS PPAQAQP-----	-----EDTFKT	108
E. coli	SFCGADAPQLRFEVG TKPVTQTPOAAVTN VAAPQAQAQTPQORA APSTRSGWNVPPA EP-----	---TTRSNNVNVKHTFDN	140
T. mar.	VVLG---NDATFEIT YEAFEPHSSYSEPLV KRAVLLTP-----	-----LNPDYTFEN	106
H. pyl.	NKVG-MHLAHSVDVR IEVAPKIQINAQSN I NYKAIKTS-----	-----VKDSYTFEN	118
P. mar.	FVVGPNSRMAHAAAM AVAESPGREENPLFI CGGVGLGKTHLMQAI	CHYRLIDPGAKVSY VSTETFTNDLIL--A IRQDRMQAERDRYR-	217
Syn. sp.	FVVGPTNRMAHAASL AVAESPGREENPLFL CGGVGLGKTHLMQAI	AHYRLEMYPNAKVY VSTERFTNDLIT--A IRQDNMEDFRSYR-	202
B. sut.	FVIGSGNRFAHAASL AVAEAPAKAYNPLFI YGGVGLGKTHLMHAI	GHYVIDHNPSAKVY LSSEKFTNEFIN--S IRDNKAVDFRNRYR-	206
M. tub.	FVIGASNRFAHAAAL ALAEAPARAYNPLFI WGESGLGKTHLLHAA	GNYAQRLFGMRVKY VSTEEFTNDFIN--S LRDDRKVAFKRSYR-	263
T. th.	SWMGPTTPWPHGGAV AVAESPGRAYNPLFI YGGRGGLGKTVLMHAV	GPLRAKRFPHMRLEY VSTETFTNDELINRPS AR-DRMTEFRERYR-	196
E. coli	FVEGKSNQLARAAAR QVADNFGGAYNPLFL YGGTGLGKTHLLHAV	GNGIMARKPNKVY MHSEFVQDMVK--A LQNNATIEEFKRYR-	227
T. mar.	FVVGPGNSFAYHAAL EVAKHPRG-YNPLFI YGGVGLGKTHLLQSI	GNVYVQNEPDLRMVY ITSEKFTNDLVD--S MKEGKLINEFREKYRK	193
H. pyl.	FVVGSCNNTVYEIAK KVAQSDTPPNPVL F YGGTGLGKTHILNAI	GNHALEK--HKKVVL VTSEDFLTDFLK--H LDNKTMDSFKAKYR-	203

FIG. 19A

P.mar. AADLLVDDIOFIEG KEYTQEEFFHTFNAL HDAGSQIVLASDRPP SQIPRLQERLMSRFS MGLIADVQAPDLETR MAILQKKAHERVGL 307  
 Syn.sp. SADFLLDIDIOFIKQ KEYTQEEFFHTFNLSL HEAGKQVVASDRAP QRIFGLQDRLISRFS MGLIADIQVPDLETR MAILQKKAEDRIRL 292  
 B.sut. NVDVLLDIDIOFLAG KEQIQEEFFHTFNVL HEESQIVISSDRPP KEIPTLEDRLRSRFS WGLITDITPPDLETR IAILRKKAKAEGLDI 296  
 M.tub. DVDVLLVDDIOFIEG KEGIQEEFFHTFNVL HNANKQIVISSDRPP KQLATLEDRLRTRFE WGLITDVQPPDLETR IAILRKKAKAOMERLAV 353  
 T.th. SVDLLVDDVQFIAG KERTQEEFFHTFNAL YEAKQIILSSDRPP KOILTLEARLSRFS WGLITDNPAPDLETR IAILKMNAS-SGPED 285  
 E.coli SVDALLIDIDIOFFAN KERSQEEFFHTFNAL LEGNQIILTSDRYP KEINGVEDRLKSRFG WGLITVAIEPPELETR VAILMKKADENDIRL 317  
 T.mar. KVDILLIDDVQFLIG KITGVQTELFTFNEL HDGKQIVICSREP QKLSEFQDRLVSRFQ MGLVAKLEPPDEETR KSIARQMLEIEHGEL 283  
 H.pyl. HCDFLLDDAQFLOG KPKLEEEFFHTFNEL HANSKQIVLISDRSP KNLAGLEDRLKSRFE WGITAKMPPDLETK LSTIVKQKCOLNQITL 293

P.mar. PRDLIQFIAGRFTSN IRELEGALTRAIAFA SITGLPMTVDSIAPM LD----PNGQGVET PKQVLDKVAEVFKVT PDEMRSASRRR-PVS 392  
 Syn.sp. PKEVIEYIASHYTSN IRELEGALIRAIAYT SLSNVAMTVENIAPV LN----PFVEKVAAA PETIITITVAQHYQLK VEELLSNSRRR-EVS 377  
 B.sut. PNEVMLYIANQIDSN IRELEGALIRVVAYS SLINKDINADLAAEA LKDII-PSSKPKVIT IKEIQRVWGQQFNIK LEDFKAKKRTK-SVA 384  
 M.tub. PDDVLELIASSIERN IRELEGALIRVTAF A SINKTPIDKALAEIV LRDLI-ADANTMQIS AATIMAATAEYFDIT VEELRGPGKTR-ALA 441  
 T.th. PEDALEYIARQVTSN IREWEGALMRASPFA SLNGVELTRAIVAACA LRHLR-P--RELEAD PLEIIRKAAGPVRPE TPGGAHGERRKEW 372  
 E.coli PGEVAFFIAKRLSN VRELEGALNRVIANA NETGRAITIDFVREA LRDL-A-LQEKLV IDNIQKTVAEYKIK VADLLSKRRSR-SVA 404  
 T.mar. PEEVLNFEVAENVDDN LRRLRGAI IKLLVYK ETIGKEVDLKEAILL LKDFIKPNRVKAMDP IDELIEIVAKVTGVP REEILSNSRV-KAL 372  
 H.pyl. PEEVMEYLAQHISDN IRQMEGAIKISVNA NIMNASIDINLAKTV LEDL--QKDHAEGSS LENIILAVAQSLNIK SSEIKVSSROK-NVA 380

P.mar. QARQVGMVLMRQGTN LSLPRIGDTFGGKDH TTVMYAIEQVEKKLS S-----DPQIA SQVQKIRDLLQIDSR RKR----- 461  
 Syn.sp. LARQVGMVLMRQHTD LSLPRIGEAFGGKDH TTVMYSCDKITQLOQ K-----DWETS QTLTSLSHRINIAGQ APES---- 447  
 B.sut. FPRQIAMVLSREMTD SSLPKIGEFGGRDH TTVTHAHEKISKLLA D-----DEQLQ QHVKEIKEQLK---- 446  
 M.tub. QSRQIAMVLCRELTD LSLPKIGQAFG-RDH TTVMYAQRKILSEMA E-----RREVF DHVKELITRIRORSK R----- 507  
 T.th. LPRQIAMVLRLETP ASLPEIGQLFGGRDH TTVRYAIQKVQELAG KP-----DREVQ GLLRITREACTDPVD NLMITCG 446  
 E.coli RPRQWAMALAKELTN HSLPEIGDAFGGRDH TTVLHACRKIEQLRE E-----SHDIK EDFSNLIRITLSS--- 467  
 T.mar. TARRIGNVYVAKNYLK SSLRTIAEKN-RSH PVVDVSVKVKVDSLL KG-----NKQLK ALIDEVIGEISRRAL SG----- 440  
 H.pyl. LARKLVVYFARLYTP NPITLSAQFLDLKDH SSISKMYSGVKQMLE EEKSPFVLSLREEIK NRLNELNDKKTAFNS SE----- 457

FIG.19B

GTGTCGCACGAGGCCGTCTGGCAACACGTTCTGGAGCA<sup>-</sup>  
 TCCGCCGCAGCATCACCGAGGTGGAGTTCCACACCTGGTT  
 TGAAAGGATCCGCCCCCTTGGGGATCCGGGACGGGGTGCTG 120  
 GAGCTCGCCGTGCCACCTCCTTTGCCCTGGACTGGATCC  
 GGCGCCACTACGCCGGCCTCATCCAGGAGGGCCCTCGGCT  
 CCTCGGGGGCCAGGCGCCCCGGT<sup>-</sup>TTGAGCTCCGGGTGGTG 240  
 CCGGGGGTCGTAGTCCAGGAGGACATCTTCCAGCCCCCGC  
 CGAGCCCCCGGCCAAGCTCAACCCGAAGATACCTTTAA  
 AACTTCGTGGTGGGGCCCAACA<sup>-</sup>ACTCCATGGCCCCACGGC 360  
 GGCGCCGTGGCCGTGGCCGAGTCCCCCGGCCGGGCCTACA  
 ACCCCCTCTTCATCTACGGGGGCCGTGGCCTGGGAAAGAC  
 CTACCTGATGCACGCCGTGGGCCCACTCCGTGCGAAGCGC 480  
 TTCCCCCACATGAGATTAGAGTACGTTTCCACGGAAACTT  
 TCACCAACGAGCTCATCAACCGGCCATCCGCGAGGGACCG  
 - GATGACGGAGTTCCGGGAGCGGTACCGCTCCGTGGACCTC 600  
 CTGCTGGTGGACGACGTCCAGTTCATCGCCGGAAAGGAGC  
 GCACCCAGGAGGAGTTTTTCCACACCTTCAACGCCCTTTA  
 CGAGGCCACAAGCAGATCATCCTCTCCTCCGACCGGCCG 720  
 CCAAGGACATCCTCACCCTGGAGGCGCGCCTGCGGAGCC  
 GCTTTGAGTGGGGCCTGATCACCGACAATCCAGCCCCCGA  
 CCTGGAAACCCGGATCGCCATCCTGAAGATGAACGCCAGC 840  
 AGCGGGCCTGAGGATCCCGAGGACGCCCTGGAGTACATCG  
 CCCGGCAGGTCACCTCCAACATCCGGGAGTGGGAAGGGGC  
 CCTCATGCGGGCATCGCCTTTCGCCTCCCTCAACGGCGTT 960  
 GAGCTGACCCGCGCCGTGGCGGCCAAGGCTCTCCGACATC  
 TTCGCCCCAGGGAGCTGGAGGCGGACCCCTTGGAGATCAT  
 CCGCAAAGCGGCGGGACCAGTTCGGCCTGAAACCCCGGGA 1080  
 GGAGCTCACGGGGAGCGCCGCAAGAAGGAGGTGGTCCTCC  
 CCGGCAGCTCGCCATGTACCTGGTGCGGGAGCTCACCCC  
 GGCTCCCTGCCCGAGATCGACCAGCTCAACGACGACCGG 1200  
 GACCACACCACGGTCCTCTACGCCATCCAGAAGGTCCAGG  
 AGCTCGCGGAAAGCGACCGGGAGGTGCAGGGCCTCCTCCG  
 CACCCTCCGGGAGGCGTGCACATGA

FIG.20A

VSHEAVWQHVLHIRRSITEVEFHTWFERIRPLGIRDGVL  
ELAVPTSFALDWIRRHVYAGLIQEGPRLPGAQAPRFELRVV  
PGVVQEDIFQPPSPPPAQAPEDTFKTSWWGPTTPWPHG 120  
GAVAVAESPGRAYNPLFIYGGRGLGKTYLMHAVGPLRAKR  
FPHMRLEYVSTETFTNELINRPSARDRMTEFRERYRSVDL  
LLVDDVQFIAGKERTQEEFFHTFNALYEAHKQIILSSDRP 240  
PKDILTLEARLRSRFEWGLITDNPAPDLETRIAILKMNAS  
SGPEDPEDALEYIARQVTSNIREWEGALMRASPFASLNGV  
ELTRAVAAKALRHLRPRELEADPLEIIRKAAGPVRPETPG 360  
GAHGERRKKEVVLPRQLAMYLVRLETPASLPEIDQLNDDR  
DHTTVLYAIQKVQELAESDREVQGLLRTLREACT

FIG.20B

ATGAACATAACGGTTCCCAAAAACTCCTCTCGGACCAGC 40  
 TTTCCCTCCTGGAGCGCATCGTCCCCTCTAGAAGCGCCAA  
 CCCCTCTACACCTACCTGGGGCTTTACGCCGAGGAAGGG 120  
 GCCTTGATCCTCTTCGGGACCAACGGGGAGGTGGACCTCG  
 AGGTCCGCCTCCCCGCCGAGGCCCAAAGCCTTCCCCGGGT 200  
 GCTCGTCCCCGCCAGCCCTTCTTCCAGCTGGTGCGGAGC  
 CTTCTGGGGACCTCGTGGCCCTCGGCCTCGCCTCGGAGC 280  
 CGGGCCAGGGGGGGCAGCTGGAGCTCTCCTCCGGGCGTTT  
 CCGCACCCGGCTCAGCCTGGCCCCTGCCGAGGGCTACCCC 360  
 GAGCTTCTGGTGCCCGAGGGGGAGGACAAGGGGGCCTTCC  
 CCTCCGGACGCGGATGCCCTCCGGGGAGCTCGTCAAGGC 440  
 CTTGACCCACGTGCGCTACGCCGCGAGCAACGAGGAGTAC  
 CGGGCCATCTTCCGCGGGGTGCAGCTGGAGTTCTCCCCC 520  
 AGGGCTTCCGGGCGGTGGCCTCCGACGGGTACCGCCTCGE  
 CCTCTACGACCTGCCCCCTGCCCCAAGGGTTCCAGGCCAAG 600  
 GCCGTGGTCCCCGCCCGGAGCGTGGACGAGATGGTGCGGG  
 TCCTGAAGGGGGCGGACGGGGCCGAGGCCGTCTCGCCCT 680  
 GGGCGAGGGGGTGTGGCCCTGGCCCTCGAGGGCGGAAGC  
 GGGGTCCGGATGGCCCTCCGCCTCATGGAAGGGGAGTTCC 760  
 CCGACTACCAGAGGGTCATCCCCCAGGAGTTCGCCCTCAA  
 GGTCCAGGTGGAGGGGGAGGCCCTCAGGGAGGCGGTGCGC 840  
 CGGGTGAGCGTCCTCTCCGACCGGCAGAACCACCGGGTGG  
 ACCTCCTTTTGGAGGAAGGCCGGATCCTCCTCTCCGCCGA 920  
 GGGGGACTACGGCAAGGGGCAGGAGGAGGTGCCCCGCCAG  
 GTGGAGGGGCGGACATGGCCGTGGCCTACAACGCCCGCT 1000  
 ACCTCCTCGAGGCCCTCGCCCCCGTGGGGGACCGGGCCCA  
 CCTGGGCATCTCCGGGCCCACGAGCCCGAGCCTCATCTGG 1080  
 GGGGACGGGGAGGGGTACCGGGCGGTGGTGGTGCCCCCTCA  
 GGGTCTAG 1128

FIG.21A

MNITVPKKLLSDQLSLLERIVPSRSANPLYTYLGLYAEEG 40  
ALILFGTNGEVDLEVRLPAEAQSLPRVLVPAQPFFQLVRS  
LPGDLVALGLASEPGQGQLELSSGRFRTRLAPAEGLYP 120  
- ELLVPEGEDKGAFPLRTRMPSELVKALTHVRYAASNEEY  
RAIFRGVQLEFSPQGFRAVASDGYRLALYDLPLPQGFQAK 200  
AVVPARSVDEMVRVLKGADGAEAVLALGEGVLALALEGGS  
GVRMALRLMEGEFPDYQRVIPQEFALKVQVEGEALREAVR 280  
RVSVLSDRQNHRVDLLLEEGRILLSAEGDYGKGQEEVPAQ  
VEGPDMAVAYNARYLLEALAPVGDRHLGISGPTSPSLIW 360  
GDGEGYRAVVVPLRVZ

FIG.21B

T.th.beta  
E.coli.bet  
P.mirab.be  
H.infl.bet  
P.put.beta  
B.cap.beta

MNITVPKKLLSDQLSLERIVPSHSANPLYTYLGLVAEEGALILFGTNGEVDLEVRPAE  
MKFTVEREHLKPLQQVSGPLGGRPTLPILGNLLQVADGTLSLTGTDLEMMVARVALV  
MKFIIEREQLLKPLQQVSGPLGGRPTLPILGNLLKVTENTLSLTGTDLEMMMARVSL  
MQFSISRENLLKPLQQVCGVLSNRENIPVANNVLLQIEDYRLTTGTDLEVELSSQTQLS  
MHFTIQREALKPLQVAGVVERQTLPVLSNVLVQVQQQLSLTGTDLEVELVGRVQLE  
MKFTIQNDILATKNLKKITRVLVKNISFPILLENILLIQVEDGTLSLTTNNLEILISKIEII  
\* \* \* \* \*

T.th.beta  
E.coli.bet  
P.mirab.be  
H.infl.bet  
P.put.beta  
B.cap.beta

AQSLP-RVLVPAQFFQLVRSPLPGDLVALGLASEPGQGQGLELSSGRFRTRLSLAPAEY  
QPHEGATTVPARKFFDICRGLP-EGAEIAVQLE---GERMLVRSGRSRFSLSTLPAADF  
QSHEIGATTVPARKFFDIWRGLP-EGAEISVELD---GDRLLVRSGRSRFSLSTLPASDF  
SSSENGTFTIPAKKFLDICRTLS-DDSEITVTFE---QDRALVQSGRSRFTLATQPAEEY  
EPAEPGEITVPARKIMDICKSLP-NDALIDIKVD---EQKLLVKAGRSRFTLSTLPANDE  
TKYIPGKTTISGRKILNICRTLS-EKSKIKMQLK---NKKMYISSENSNYILSTLSADTF  
\* \* \* \* \*

T.th.beta  
E.coli.bet  
P.mirab.be  
H.infl.bet  
P.put.beta  
B.cap.beta

PELLVPEGEDKGAFFLTRMPSEGLVKALTHVRYAASNEEYRAIFRGVQLEFSPQFRAV  
PNLDD--WQSEVEFTLPQAT-----MKRLIEATQFSMAHQDVRYVYINGMLFETEGEELRTV  
PNLDD--WQSEVEFTLPQAT-----LKRLIESTQFSMAHQDVRYVYINGMLFETENTELRTV  
PNLTD--WQSEVDLFELPONT-----LRRLIEATQFSMANQDARYFINGMKFETEGNLLRTV  
PTVEE--GPGSLTCNLEQSK-----LRRLIERTSFAMAQQDVRYVYINGMLLEVSRNTLRAV  
PNHQN--FDYISKFDISSNI-----LKEMIEKTEFSMGKQDVRYVYINGMLLEKKDKFLRSV  
\* \* \* \* \*

T.th.beta  
E.coli.bet  
P.mirab.be  
H.infl.bet  
P.put.beta  
B.cap.beta

ASDGYRLALYDLPLPQGFQA--KAVVPARSVDEMVRVLKGADGAELVLAGGVLALE  
ATDGHRLAVCSMPIGQSLPS-HSVIVPRKGVIELMRMLDG-GDNPLRVQIGSNNIRAHVG  
ATDGHRLAVCANDIGQSLPG-HSVIVPRKGVIELMRLLDGSGESLQIQIGSNNLRAHVG  
ATDGHRLAVCTISLEQELQN-HSVILPRKGVLELVRLLET-NDEPARLQIGTNNLRVHIK  
STDGHRLALCSMSAPIEQEDRHQVIVPRKGILELARLLTD-PEGMVSIVLQHHIRATTG  
ATDGYRLAISYTQLKIDINF-FSIIIPNKAVMELLKLLNT-OPQLNLLIGSNSIRIYTK  
\*\* \*\*\* \*  
\* \* \* \* \*

FIG.22A



T. th. beta	GGSGVRNALRLMEGEFPDYQRVIPQEFALKVQVEGEALREAVRRVSVLSDRQNRHVDLLL
E. coli. bet	---DFIFTSKLVDRFPDYRRVLPKNIPDKHLEAGCDLLKQAFARAAILSNEKFRGVRLLYV
P. mirab. be	---DFIFTSKLVDRFPDYRRVLPKNIPDKHLEAGCDLLKQAFSRAAILSNEKFRGVRINL
H. infl. bet	---NTVFTSKLIDGRFPDYRRVLPKNATKIVBGNWEMLKQAFARASILSNERARSVRLSL
P. put. beta	---EFIFTSKLVDRFPDYRRVLPKNATKIVBGNWEMLKQAFARASILSNEKYGIRLQL
B. cap. beta	---NLIFTTQLIEGEYFDYKSVLFKEKNPIITNSILLKSLLRVAIIAHEKFCGIEIKI

T. th. beta	EEGRILLSAEGDYGK-GQEEVPAQVEGPDMAVAYNARYLLEALAPVG-DRAHLGISGPTS
E. coli. bet	SENQLKITANNPEQEEAAEILLDTYSGAEMEIGFNVSYLLDVNLALNCENVRMMLTDSVS
P. mirab. be	TNGQLKITANNPEQEEAAEIVDVQYQGEEMEIGFNVSYLLDVNLALNCENVRMMLTDSVS
H. infl. bet	KENQLKITASNTHEHEAAEIVDVNNGEELVGFNVTYLLDVNLALNCENVRMMLTDAFS
P. put. beta	AAGQLKIQANNPEQEEAAEISVDYEGSSLEIGFNVSYLLDVNLALNCENVRMMLTDAFS
B. cap. beta	ENGKFKVLSDNQEEETAEDLFEIDYFGEKIEISINVVYLLDVNNIKSENIALFLNKSKS

T. th. beta	PSLIWGDG-EGYRAVWVPLRVZ	(ID#108)
E. coli. bet	SVQIEDAASQSAAYVVMFMRLLZ	(ID#109)
P. mirab. be	SVQVENVASAAAAYVVMFMRLL-	(ID#110)
H. infl. bet	SCLIENCESSCEYVIMFMRLL-	(ID#111)
P. put. beta	SALLQEAAGNDSSYVVMFMRLL-	(ID#112)
B. cap. beta	SIQIEAENNSSNAYVVMFLKR-	(ID#113)

FIG. 22B

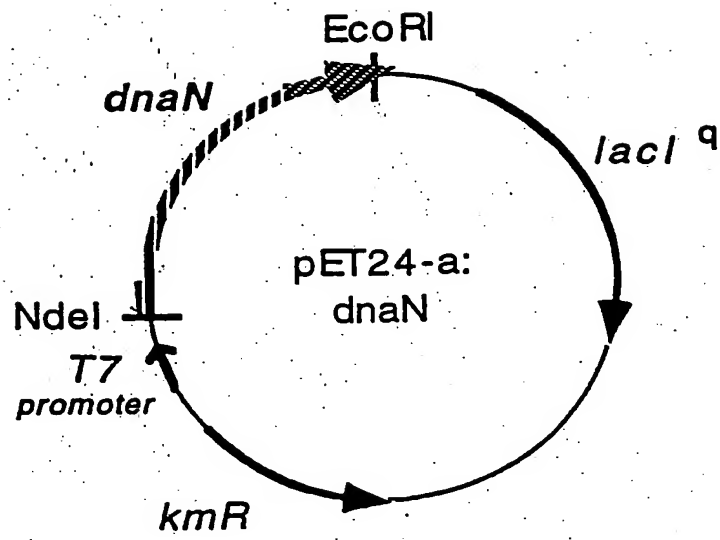
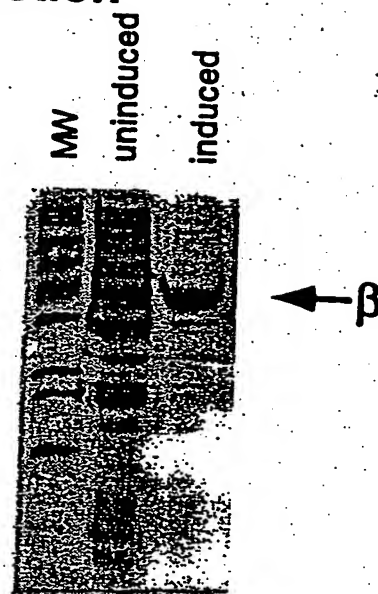


FIG.23

**FIG.24A** Induction



Lysis

Heat Step

**FIG.24B** MonoQ Column

Fraction: 5 7 9 11 13 15 17 19 21 23 25

$\beta$  →

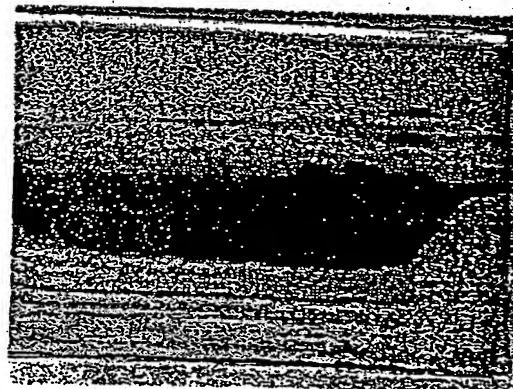


FIG.25A

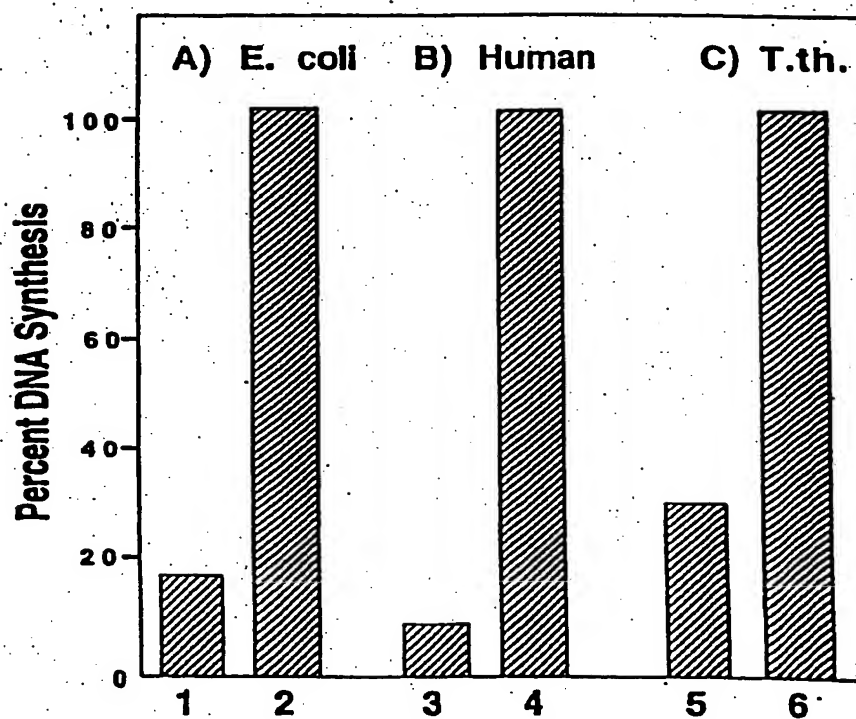
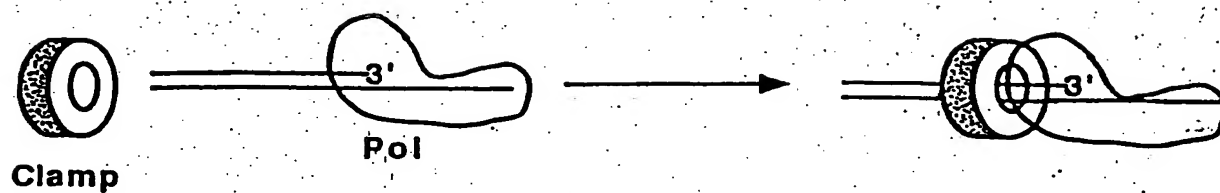


FIG.25B

FIG.26A

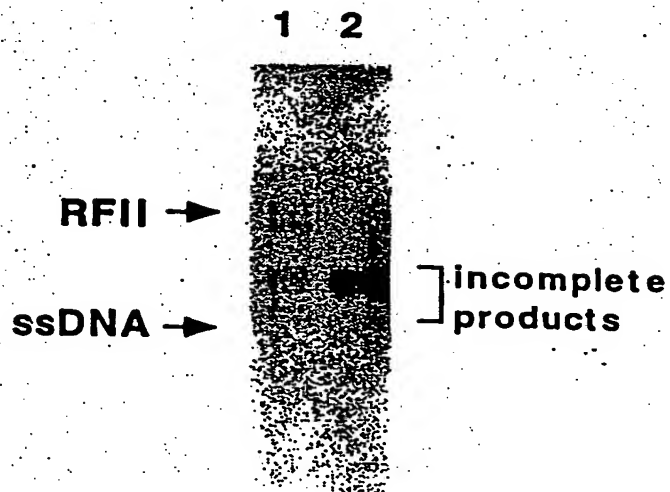
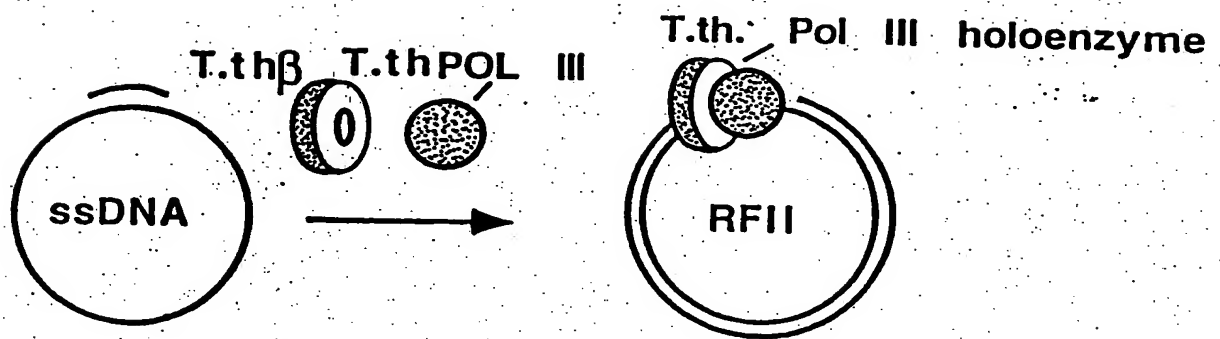


FIG.26B

$\alpha$        $\tau$        $\delta$        $\delta'$  SSB  $\epsilon$



FIG. 27

2 4 6 8 10 12 14 16 18 20 22 24 26 28 sds 30 32 34 36 38 40 42 44 46 48 50 52 54



$\tau$   
 $\delta'$   
 $\delta$

FIG. 28

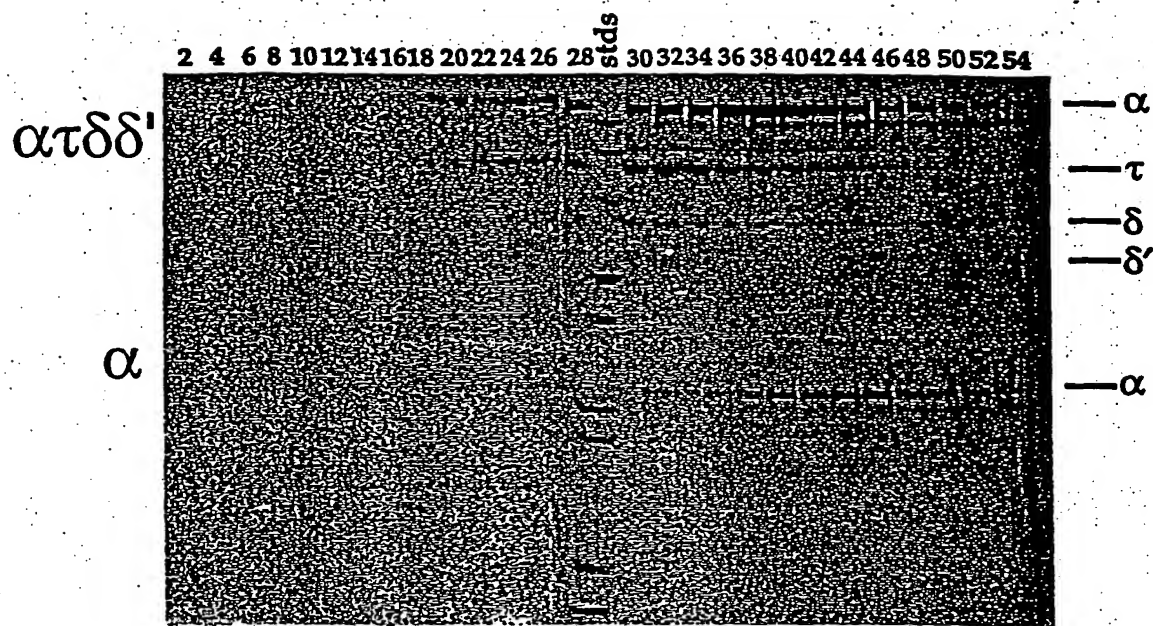


FIG. 29

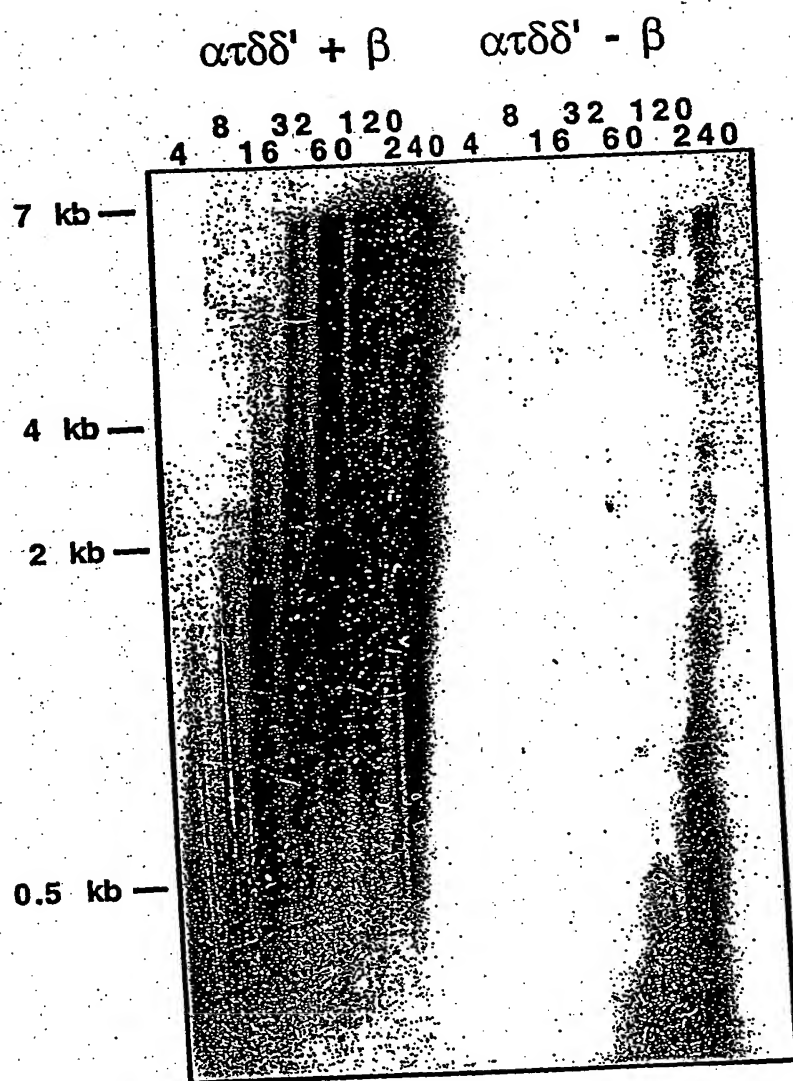
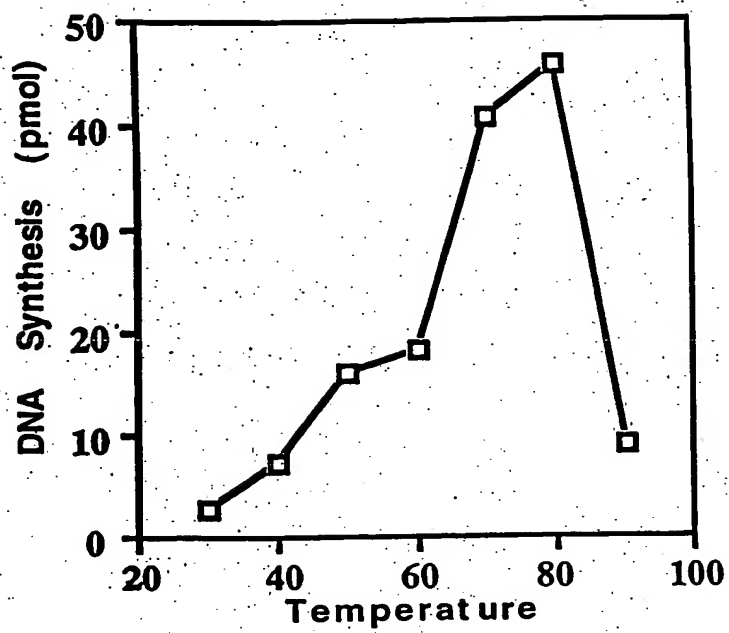
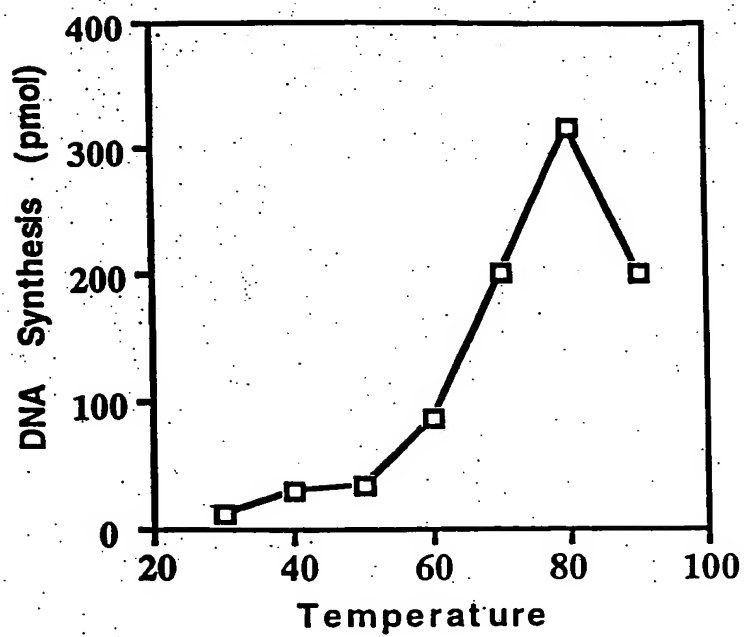


FIG. 30





**FIG. 31**



**FIG. 32**

FIG. 33A

FIG. 33B

FIG. 33C

FIG. 33D

FIG. 33E

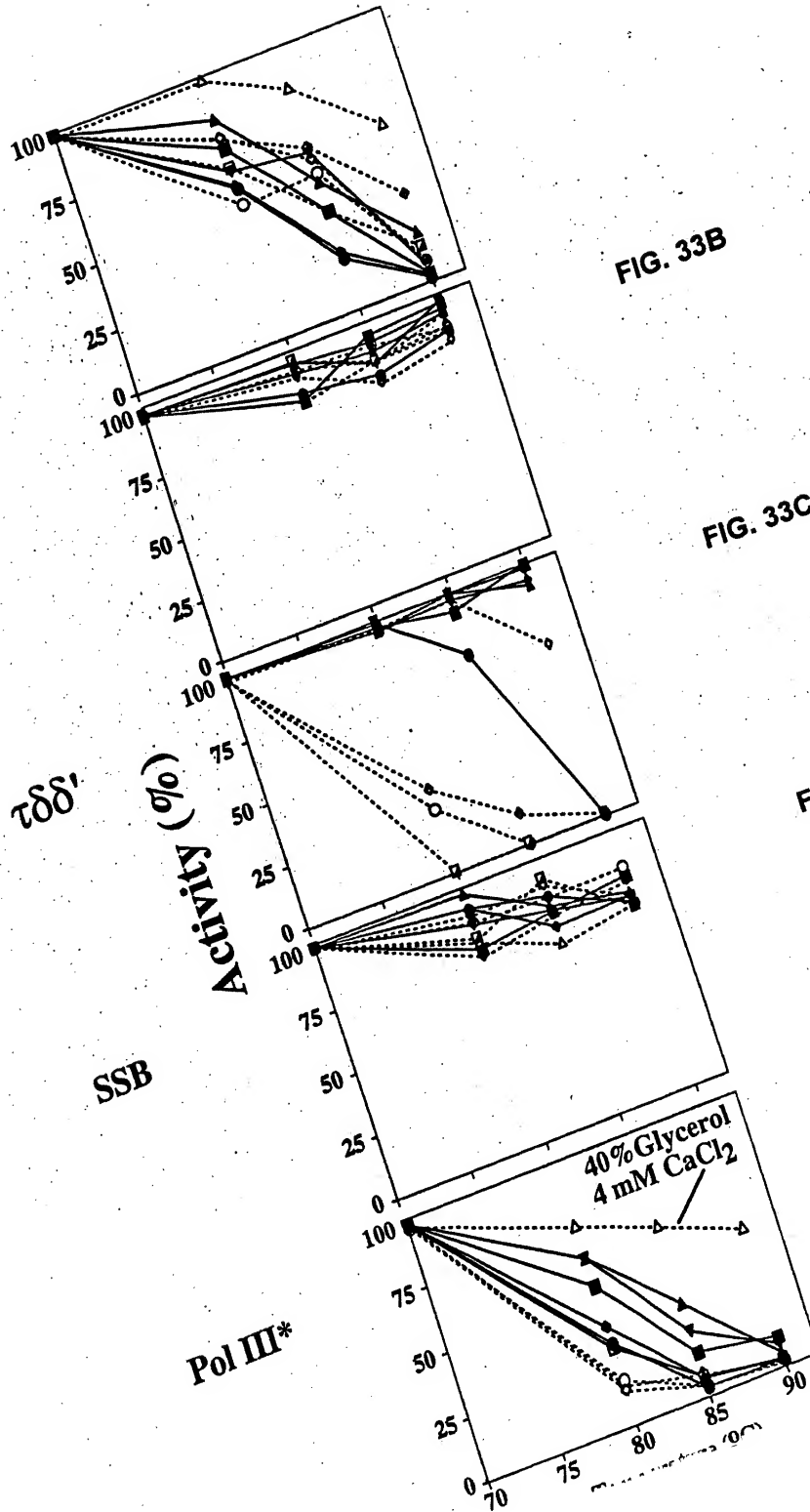
$\alpha$

$\beta$

$\tau\delta\delta'$

SSB

Pol III\*



ATGAGTAAGGATTTTCGTCCACCTTCACCTGCACACCCAGTTCTCACTCCT	
GGACGGGGCTATAAAGATAGACGAGCTCGTGAAAAAGGCAAAGGAGTATG	100
GATACAAAGCTGTCGGAATGTCAGACCACGGAAACCTCTTCGGTTTCGTAT	
AAATTCTACAAAGCCCTGAAGGCGGAAGGAATTAAGCCCATAATCGGCAT	200
GGAAGCCTACTTTACCACGGGTTTCGAGGTTTGACAGAAAGACTAAAACGA	
GCGAGGACAACATAACCGACAAGTACAACCACCACCTCATACTTATAGCA	300
AAGGACGAAAAGGTCTAAAGAAGTTAATGAAGCTCTCAACCCCTCGCCTAC	
AAAGAAGGTTTTTTACTACAAACCCAGAATTGATTACGAACTCCTTGAAAA	400
GTACGGGGAGGGCCTAATAGCCCTTACCGCATGCCTGAAAGGTGTTCCCA	
CCTACTACGCTTCTATAAACGAAGTGAAAAAGGCGGAGGAATGGGTAAAG	500
AAGTTCGAAGGATATATTCGGAGATGACCTTTATTTAGAACTTCAAGCGAA	
CAACATTCCAGAACAGGAAGTGGCAAACAGGAACTTAATAGAGATAGCCA	600
AAAAGTACGATGTGAACTCATAGCGACGCAGGACGCCCACTACCTCAAT	
CCCGAAGACAGGTACGCCACACGGTTCTTATGGCACTTCAAATGAAAAA	700
GACCATTACGAACTGAGTTCGGGAACTTCAAGTGTTCAAACGAAGACC	
TTCACTTTGCTCCACCCGAGTACATGTGGAAAAAGTTTGAAGGTAAGTTC	800
GAAGGCTGGGAAAAGGCACTCCTGAACACTCTCGAGGTAATGGAAAAGAC	
AGCGGACAGCTTTGAGATATTTGAAAACCTCACCTACCTCCTTCCCAAGT	900
ACGACGTTCCGCCCGACAAAACCTTGAGGAATACTCAGAGAACTCGCG	
TACAAAGGTTTAAAGACAGAGGATAGAAAGGGGACAAGCTAAGGATACTAA	1000
AGAGTACTGGGAGAGGCTCGAGTACGAACTGGAAGTTATAAACAAAATGG	
GCTTTGCGGGATACTTCTTGATAGTTTCAGGACTTCATAAACTGGGCTAAG	1100
AAAAACGACATACCTGTTGGACCCGGAAGGGGAAGTGCTGGAGGTTCCCT	
CGTCGCATACGCCATCGGAATAACGGACGTTGACCCTATAAAGCACGGAT	1200
TCCTTTTTGAGAGGTTCTTAAACCCCGAAAGGGTTTCCATGCCGGATATA	
GACGTGGATTTCTGTGAGGACAACAGGGAAAAGGTCATAGAGTACGTAAG	1300
GAACAAGTACGGACACGACAACGTAGCTCAGATAATCACCTACAACGTAA	
TGAAGGCGAAGCAAACACTGAGAGACGTCGCAAGGGCCATGGGACTCCCC	1400
TACTCCACCGCGGACAAACTCGCAAAACTCATTCCTCAGGGGGACGTTCA	
GGGAACGTGGCTCAGTCTGGAAGAGATGTACAAAACGCCTGTGGAGGAAC	1500
TCCTTCAGAAGTACGGAGAACAAGAACGACATAGAGGACAACGTAAAG	
AAGTTCAGACAGATATGCGAAGAAAGTCCGGAGATAAAACAGCTCGTTGA	1600
GACGGCCCTGAAGCTTGAAGGTCTCACGAGACACACCTCCCTCCACGCCG	
CGGGAGTGGTTATAGCACCAAAGCCCTTGAGCGAGCTCGTTCCCCTCTAC	1700
TACGATAAAGAGGGCGAAGTCGCAACCCAGTACGACATGGTTCAGCTCGA	
AGAACTCGGTCTCCTGAAGATGGACTTCCTCGGACTCAAACCCCTCACAG	1800
AACTGAACTCATGAAAGAACTCATAAAGGAAAGACACGGAGTGGATATA	
AACTTCCTTGAACCTTCCCCTTGACGACCCGAAAGTTTACAAACTCCTTCA	1900
GGAAGGAAAAACACGGGAGTGTTCCAGCTCGAAAGCAGGGGAATGAAAG	
AACTCCTGAAGAACTAAAGCCCGACAGCTTTGACGACATCGTTGCGGTC	2000
CTCGCACTCTACAGACCCGGACCTCTAAAGAGCGGACTCGTTGACACATA	
CATTAAGAGAAAGCACGGAAAAGAACCCGTTGAGTACCCCTTCCCAGGAGC	2100
TTGAACCCGTCCTTAAGGAAACCTACGGAGTAATCGTTTATCAGGAACAG	
GTGATGAAGATGTCTCAGATACTTTCCGGCTTTACTCCCAGGAGAGGCGGA	2200
TACCCCTCAGAAAGGCGATAGGTAAGAAGAAAGCGGATTTAATGGCTCAGA	
TGAAGACAAGTTTCATACAGGGAGCGGTGGAAAGGGGATACCCTGAAGAA	2300
AAGATAAGGAAGCTCTGGGAAGACATAGAGAAGTTCGCTTCTTACTCCTT	
CAACAAGTCTCACTCGGTAGCTTACGGGTACATCTCCTACTGGACCGCCT	2400

FIG. 34A

ACGTTAAAGCCCACTATCCCGCGGAGTTCTTCGCGGTAAAACTCACAACT	
GAAAAGAACGACAACAAGTTCCTCAACCTCATAAAAGACGCTAAACTCTT	2500
CGGATTTGAGATACTTCCCCCGACATAAACAAGAGTGATGTAGGATTTA	
CGATAGAAGGTGAAAACAGGATAAGGTTCTGGGCTTGCGAGGATAAAGGGA	2600
GTGGGAGAGGAACTGCTAAGATAATCGTTGAAGCTAGAAAGAAGTATAA	
GCAGTTCAAAGGGCTTGCGGACTTCATAAACAAAACCAAGAACAGGAAGA	2700
TAAACAAGAAAGTCGTGGAAGCACTCGTAAAGGCAGGGGCTTTTGACTTT	
ACTAAGAAAAAGAGGAAAGAACTACTCGCTAAAGTGGCAAACCTCTGAAAA	2800
AGCATTAATGGCTACACAAAACCTCCCTTTTCGGTGCACCGAAAGAAGAAG	
TGGAAGAACTCGACCCCTTAAAGCTTGAAAAGGAAGTTCTCGGTTTTTAC	2900
ATTTCAGGGCACCCCTTGACAACCTACGAAAAGCTCCTCAAGAACCGCTA	
CACACCCATTGAAGATTTAGAAGAGTGGGACAAGGAAAGCGAAGCGGTGC	3000
TTACAGGAGTTATCACGGAACCTCAAAGTAAAAAAGACGAAAACCGGAGAT	
TACATGGCGGTCTTCAACCTCGTTGACAAGACGGGACTAATAGAGTGTGT	3100
CGTCTTCCCGGGAGTTTACGAAGAGGCAAAGGAACTGATAGAAGAGGACA	
GAGTAGTGGTAGTCAAAGGTTTTCTGGACGAGGACCTTGAAACGGAAAAT	3200
GTCAAGTTCGTGGTGAAAGAGGTTTTCTCCCCTGAGGAGTTCGCAAAGGA	
GATGAGGAATACCCTTTATATATTCTTAAAAAGAGAGCAAGCCCTAAACG	3300
GCGTTGCCGAAAACTAAAGGGAATTATTGAAAACAACAGGACGGAGGAC	
GGATACAACCTTGGTTCTCACGGTTGATCTGGGAGACTACTTCGTTGATTT	3400
AGCACTCCCAAGATATGAACTAAAGGCTGACAGAAAGGTTGTAGAGG	
AGATAGAAAACTGGGAGTGAAGGTCATAATTTAGTAAATAACCCTTACT	3500
TCCGAGTAGTCCCC	

**FIG. 34B**

MSKDFVHLHLHTQFSLLDGAIKIDELVKKAKEYGYKAVGMSDHGNLFGSY	
KFYKALKAEGIKPIIGMEAYFTTGSRFDRKTKTSEDNITDKYNHHLILIA	100
KDDKGLKNLMKLSTLAYKEGFYKPRIDYELLEKYGEGLIALTACLKGV	
TYYASINEVKKAEWVKKFKDIFGDDLYLELQANNIPEQEVANRNILIEIA	200
KKYDVKLIATQDAHYLNPEDRYAHTVLMALQMKKTIHELSSGNFKCSNED	
LHFAPPEYMWKKFEGKFEGWEKALLNTLEVMEKTADSFEIFENSTYLLPK	300
YDVPPDKTLEEYLRELAYKGLRQRIERGQAKDTKEYWERLEYELEVINKM	
GFAGYFLIVQDFINWAKKNDIPVGPGRGSAGGSLVAYAIGITDVPDIKHG	400
FLFERFLNPERVSMPCDIDVDFCQDNREKVIEWVRNKYGHDNVAQIITYNV	
MKAKQTLRDVARAMGLPYSTADKLAKLIPQGDVQGTWLSLEEMYKTPVEE	500
LLQKYGEHRDIEDNVKKFRQICEESPEIKQLVETALKLEGLTRHTSLHA	
AGVVIAPKPLSELVPLYDKEGEVATQYDMVQLEELGLLKMDFLGLKTLT	600
ELKLMKELIKERHGVDFINLELPLDDPKVYKLLQEGKTTGVFQLESRGMK	
ELLKKLKPDSDDDIVAVLALYRPGPLKSGLVDTYIKRKHGKEPVEYPFPE	700
LEPVLKETYGIVYQEQVMKMSQILSGFTPGEADTLRKAIGKKKADLMAQ	
MKDKFIQGAVERGYPEEKIRKLWEDIEKFASYSFNKSHSVAYGYISYWTA	800
YVKAHYPAEFFAVKLTTEKNDNKFLNLIKDAKLFGEILPPDINKSDVGF	
TIEGENRIRFGLARIKGVGEETAKIIVEARKKYKQFKGLADFINKTKNRK	900
INKKVVEALVKAGAFDFTKKKRKELLAKVANSEKALMATQNSLFGAPKEE	
VEELDPLKLEKEVLGFYISGHPLDNYEKLLKNRYTPIEDLEEWKKESEAV	1000
LTGVITELKVKKTKNGDYMVFNLVDKTGLIECVVFPVGYEEAKELIEED	
RVVVVKGFLDEDLETENVKFVVKVFSPEEFKEMRNTLYIFLKREQALN	1100
GVAEKLKGIENNRTEDGYNLVLTVDLGDYFVDLALPQDMKCLKADRKVVE	
EIEKLGVKVII	1161

FIG. 35

ATGAACTACGTTCCCTTCGCGAGAAAGTACAGACCGAAATTCTTCAGGGA	
AGTAATAGGACAGGAAGCTCCCGTAAGGATACTCAAAAACGCTATAAAAA	100
ACGACAGAGTGGCTCACGCCTACCTCTTTGCCGGACCGAGGGGGGTTGGG	
AAGACGACTATTGCAAGAATTCTCGCAAAAGCTTTGAACTGTAAAAATCC	200
CTCCAAAGGTGAGCCCTGCGGTGAGTGCGAAAACCTGCAGGGAGATAGACA	
GGGGTGTGTTCCCTGACTTAATTGAAATGGATGCCGCCTCAAACAGGGGT	300
ATAGACGACGTAAGGGCATTAAAAGAAGCGGTCAATTACAAACCTATAAA	
AGGAAAGTACAAGGTTTACATAATAGACGAAGCTCACATGCTCACGAAAG	400
AAGCTTTCAACGCTCTCTTAAAAACCCTCGAAGAGCCCCCTCCAGAACT	
GTTTTCGTCCTTTGTACCACGGAGTACGACAAAATTCTTCCCACGATACT	500
CTCAAGGTGTCAGAGGATAATCTTCTCAAAGGTAAGAAAGGAAAAAGTAA	
TAGAGTATCTAAAAAGATATGTGAAAAGGAAGGGATTGAGTGCGAAGAG	600
GGAGCCCTTGAGGTTCTGGCTCATGCCTCTGAAGGGTGTCATGAGGGATGC	
AGCCTCTCTCCTGGACCAGGCGAGCGTTTACGGGGAAGGCAGGGTAACAA	700
AAGAAGTAGTGGAGAAGCTTCTCGGAATTCTCAGTCAGGAAAGCGTTAGG	
AGTTTTCTGAAATTGCTTCTGAACTCAGAAGTGGACGAAGCTATAAAGTT	800
CCTCAGAGAACTCTCAGAAAAGGGCTACAACCTGACCAAGTTTTGGGAGA	
TGTTAGAAGAGGAAGTGAGAAACGCAATTTTAGTAAAGAGCCTGAAAAAT	900
CCCGAAAGCGTGTTTCAAGACTGGCAGGATTACGAAGACTTCAAAGACTA	
CCCTCTGGAAGCCCTCCTCTACGTTGAGAACCTGATAAACAGGGGTAAAG	1000
TTGAAGCGAGAACGAGAGAACCTTAAGAGCCTTTGAACTCGCGGTAATA	
AAGAGCCTTATAGTCAAAGACATAATTCCCGTATCCCAGCTCGGAAGTGT	1100
GGTAAAGGAAACCAAAAAGGAAGAAAAGAAAGTTGAAGTAAAAGAAGAGC	
CAAAAGTAAAAGAAGAAAAACCAAAGGAGCAGGAAGAGGACAGGTTCCAG	1200
AAAGTTTTTAAACGCTGTGGACGGCAAAATCCTTAAAAGAATACTTGAAGG	
GGCAAAAAGGGAAGAAAGAGACGGAAAAATCGTCCTAAAGATAGAAGCCT	1300
CTTATCTGAGAACCATGAAAAAGGAATTTGACTCACTAAAGGAGACTTTT	
CCTTTTTTTAGAGTTTGAACCCGTGGAGGATAAAAAAAACCTCAGAAGTC	1400
CAGCGGGACGAGGCTGTTTTAAAGGTAAAGGAGCTCTTCAATGCAAAAAT	
ACTCAAAGTACGAAGTAAAAGCTAAGGTCATAAAGGTGAGAATGCCCGTG	1500
GAAGAGATAGGGCTGTTTAACGCACTAATAGACGGCTTGCCCAGGTACGC	
ACTCACGAGGACGAAGGAAAAGGGAAAGGGAGAAGTTTTTCGTTTTAGCGA	1600
CTCCTTATAAAGTCAAGGAATTGATGGAAGCTATGGAGGGTATGAAAAAA	
CACATAAAGGATTTAGAAATCCTCGGAGAGACGGATGAGGATTTAACTTT	1700
TTAAAGTATGGGTGTATCTGAGCAAAGGTTTAAGCTAAAAACAAACCTGA	
AACCCGCAGGGGACCAGCCGAAAGCCATAAAAAAACTCCTTGAAAACCTA	1800
AGGAAAGGCGTAAAAGAACAACACTTCTCGGAGTCACGGGAAGCGGAAA	
GACTTTTTACTCTAGCAAACGTAATAGCGAAGTACAACAAACCAACTCTTG	1900
TGGTAGTTCACAACAAAATTCTCGCGGCACAGCTATACAGGGAGTTTAAA	
GAAGTATTCCTGAAAACGCTGTAGAGTACTTTGTCTTTACTACGACTA	2000
TTACCAACCTGAAGCCTACATTCCCGAAAAAGATTTATACATAGAAAAGG	
ACGCGAGTATAAACGAAAGCTGGAACGTTTCAGACACTCCGCCACGATAT	2100
CCGTTCTAGAAAGGAGGGACGTTATAGTAGTTGCTTCAGTTTCTTGCCATA	
TACGGACTCGGGAAACCTGAGCACTACGAAAACCTGAGGATAAACTCCA	2200
AAGGGGAATAAGACTGAACTTGAGTAAGCTCCTGAGGAAACTCGTTGAGC	
TAGGATATCAGAGAAATGACTTTGCCATAAAGAGGGCTACCTTCTCGGTT	2300
AGGGGAGACGTGGTTGAGATAGTCCCTTCTCACACGGAAGATTACCTCGT	
GAGGGTAGAGTTCTGGGACGACGAAGTTGAAAGAATAGTCCTCATGGACG	2400
CTCTGAAC	

FIG. 36

MNYVPFARKYRPKFFREVIGQEAPVRILKNAIKNDRVAHAYLFAGPRGVG	
KTTIARILAKALNCKNPSKGEPCGECENCREIDRGVFPDLIEMDAASNRG	100
IDDVRALKEAVNYKPIKGKYKVYIIIDEAHMLTKEAFNALLKTLEPPPR	
VFVLCTTEYDKILPTILSRCQRIIFSQRKEKVEYLLKKICEKEGIECEE	200
GALEVLAHASEGCMRDAASLLDQASVYGEGRVTKEVVENFLGILSQESVR	
SFLKLLLNSEVDEAIKFLRELSEKGYNLTKFWEMLEEEVRNAILVKSLKN	300
PESVVQNWQDYEDFKDYPLEALLYVENLINRGKVEARTREPLRAFELAVI	
KSLIVKDIIPVSQLGSAVKETKKEEKKEVKEEPKVKEEKPKKEQEDRFQ	400
KVLNAVDGKILKRILEGAKREERDGIIVLKIEASYLRMTMKKEFDLSKETF	
PFLEFEPVEDKKKPKSSGTRLF	473

FIG. 37

ATGCGCGTTAAGGTGGACAGGGAGGAGCTTGAAGAGGTTCTTAAAAAGC	
AAGAGAAAGCACGGAAAAAAGCCGCACTCCCGATACTCGCGAACTTCT	100
TACTCTCCGCAAAGAGGAAAACCTTAATCGTAAGGGCAACGGACTTGAA	
AACTACCTTGTAAGTCTCCGTAAAGGGGAGGTTGAAGAGGAAGGAGAGGT	200
TTGCGTCCACTCTCAAAAACCTCTACGATATAGTCAAGAACTTAAATCCG	
CTTACGTTTACCTTCATACGGAAGGTGAAAACTCGTCATAACGGGAGGA	300
AAGAGTACGTACAACTTCCGACAGCTCCCGCGGAGGACTTTCCCGAATT	
TCCAGAAATCGTAGAAGGAGGAGAAACACTTTCGGGAAACCTTCTCGTTA	400
ACGGAATAGAAAAGGTAGAGTACGCCATAGCGAAGGAAGAAGCGAACATA	
GCCCTTCAGGGAATGTATCTGAGAGGATACGAGGACAGAATTCACTTTGT	500
GTTCCGGACGGTCACAGGCTTGCACTTTATGAACCTCTACGTAAACATTGA	
AAAGAGTGAAGACGAGTCTTTTGCTTACTTCTCCACTCCCGAGTGGAAC	600
TCGCCGTTAGCTCCTGGAAGGAGAATTCCCGGACTACATGAGTGTATCC	
CTGAGGAGTTTTCGGCGGAAGTCTTGTTGAGACAGAGGAAGTCTTAAAG	700
GTTTTAAAGAGGTTGAAGGCTTTAAGCGAAGGAAAAGTTTTTCCCGTGAA	
GATTACCTTAAGCGAAAACCTTGCCATCTTTGAGTTCGCGGATCCGGAGT	800
TCGGAGAAGCGAGAGAGGAAATTGAAGTGGAGTACACGGGAGAGCCCTTT	
GAGATAGGATTCAACGGAAATACCTTATGGAGGCGCTTGACGCCTACGAC	900
AGCGAAAGAGTGTGGTTCAAGTTCACAACCCCCGACACGGCCACTTTATT	
GGAGGCTGAAGATTACGAAAAGGAACCTTACAAGTGCATAATAATGCCGA	1000
TGAGGGTGTAGCCATGAAAAAGCTTTAATCTTTTTATTGAGCTTGAGCC	
TTTTAATTCTGCGTTTAGCGAAGCCAAACCCAAGTCTTC	1090

FIG. 38

MRVKVDREELEEVLLKARESTEKKAALPILANFLLSAKEENLIVRATDLE	
NYLVVSVKGEVEEEGEVVCVHSQKLYDIVKNLNSAYVYLHTEGEKLVITGG	100
KSTYKLPTAPAEDFPEFPEIVEGGETLSGNLLVNGIEKVEYAIKEEANI	
ALQGMYLGRGYEDRIHFVGS DGHRLALYEPLGEFSKELLI PRKSLKVLKKL	200
ITGIEDVNIKSEDESFA YFSTPEWKLA VRLLEGEFPDYMSVIP EEFSAE	
VL FETEEVLKVLKRLKALSEGKVFPVKITLSENLAIFE FADPEFGEAREE	300
IEVEYTGE PF EIGFNGKYLMEALDAYDSERVWFKFTTPDTATLLEAEDYE	
KEPYKCIIMPMRV	363

FIG. 39



GTGGAAACCACAATATTCCAGTTCAGAAAACCTTTTTTCACAAAACCTCC	
GAAGGAGAGGGTCTTCGTCCTTCATGGAGAAGAGCAGTATCTCATAAGAA	100
CCTTTTTGTCTAAGCTGAAGGAAAAGTACGGGGAGAATTACACGGTTCTG	
TGGGGGGATGAGATAAGCGAGGAGGAATTCTACACTGCCCTTCCGAGAC	200
CAGTATATTTCGGCGGTTCAAAGGAAAAAGCGGTGGTCATTTACAACCTTCG	
GGGATTTCTGAAGAAGCTCGGAAGGAAGAAAAAGGAAAAAGAAAGGCTT	300
ATAAAAGTCCTCAGAAACGTAAAGAGTAACTACGTATTTATAGTGTACGA	
TGCGAAACTCCAGAAACAGGAACTTCTTCGGAACCTCTGAAATCCGTAG	400
CGTCTTTCGGCGGTATAGTGGTAGCAAACAGGCTGAGCAAGGAGAGGATA	
AAACAGCTCGTCCCTTAAGAAGTTCAAAGAAAAAGGGATAAACGTAGAAAA	500
CGATGCCCTTGAATACCTTCTCCAGCTCACGGGTTACAACCTTGATGGAGC	
TCAAACCTTGAGGTTGAAAAACTGATAGATTACGCAAGTGAAAAGAAAATT	600
TTAACACTCGATGAGGTAAAGAGAGTAGCCTTCTCAGTCTCAGAAAACGT	
AAACGTATTTGAGTTCGTTGATTTACTCCTCTTAAAAGATTACGAAAAGG	700
CTCTTAAAGTTTTTGACTCCCTCATTTCTTCGGAATACACCCCTCCAG	
ATTATGAAAATCCTGTCTCCTATGCTCTAAAACCTTACACCCTCAAGAG	800
GCTTGAAGAGAAGGGAGAGGACCTGAATAAGGCGATGGAAAGCGTGGGAA	
TAAAGAACAACCTTCTCAAGATGAAGTTCAAATCTTACTTAAAGGCAAAC	900
TCTAAAGAGGACTTGAAGAACCTAATCCTCTCCCTCCAGAGGATAGACGC	
TTTTTCTAAACCTTACTTTTCAGGACACAGTGCAGTTGCTGGGGATTTCTT	1000
GACCTCAAGACTGGAGAGGGAAGTTGTGAAAAATACTTCTCATGGTGGAT	
AATCTTTTTTATGAAGTTTGCGGTTTGCGTTTTTCCCGGTTCT	1093

FIG. 40

VETTIFQFQKTFFTKPPKERVFLHGEEQYLIRTFLSKLKEKYGENYTVL	
WGDEISEEEFYTALSETSI FGGSKEKAVVIYNFGDFLKKLGRKKKEKERL	100
IKVLRNVKSNIYVFIVYDAKLQKQELSSEPLKSVASFGGIVVANRLSKERI	
KQLVLKKFKEKGINVENDALEYLLQLTGYNLMELKLEVEKLIDYASEKKI	200
LTLDEVKRVAFSVSENVNVFEFVDLLLLLDYKALKVLDLSISFGIHPLQ	
IMKILSSYALKLYTLKRLEEKGEDLNKAMESVGIKNNFLKMKFKSYLKAN	300
SKEDLKNLILSLQRIDAFSKLYFQDTVQLLRDFLTSLRLEREVVKNNTSHGG	

FIG. 41

ATGGAAAAAGTTTTTTTGGAAAAACTCCAGAAAACCTTGACATACCCGG	
AGGACTCCTTTTTTACGGCAAAGAAGGAAGCGGAAAGACGAAAAACAGCTT	100
TTGAATTTGCAAAGGTATTTTATGTAAGGAAAACGTACCTGGGGATGCG	
GAAGTTGTCCCTCCTGCAAACACGTAAACGAGCTGGAGGAAGCCTTCTTT	200
AAAGGAGAAATAGAAGACTTTAAAGTTTATAAGACAAGGACGGTAAAAAG	
CACTTCGTTTACCTTATGGGCGAACATCCCGACTTTGTGGTAATAATCCC	300
GAGCGGACATTACATAAAGATAGAACAGATAAGGGAAGTTAAGAACTTTG	
CCTATGTGAAGCCCGCACTAAGCAGGAGAAAAGTAATTATAATAGACGAC	400
GCCACGCGATGACCTCTCAGGCGGCAAACGCTCTTTTAAAGGTATTGGA	
AGAGCCACCTGCGGACACCACCTTTATCTTGACCACGAACAGGCGTTCTG	500
CAATCCTGCCGACTATCCTCTCCAGAACTTTTCAAGTGGAGTTCAAGGGC	
TTTTCAGTAAAAGAGGTTATGGAAATAGCGAAAGTAGACGAGGAAATAGC	600
GAAACTCTCTGGAGGCAGTCTAAAAAGGGCTATCTTACTAAAGGAAAACA	
AAGATATCCTAAACAAAGTAAAGGAATTCTTGAAAACGAGCCGTAAAAA	700
GTTTACAAGCTTGCAAGTGAATTCGAAAAGTGGGAACCTGAAAAGCAAAA	
ACTCTTCCTTGAAATTATGGAAGAATTGGTATCTCAAAAATTGACCGAAG	800
AGAAAAAAGACAATTACACCTACCTTCTTGATACGATCAGACTCTTTAA	
GACGGACTCGCAAGGGGTGTAAACGAACCTCTGTGGCTGTTTACGTTAGC	900
CGTTCAGGCGGATTAATAAACCGTTATTGATTCCGTAACATTTAAACCTT	
AATCTAAATTATGAGAGCCTTTGAAGGAGGTCTGGTATGGAAAATTTGAA	1000
GATTAGATATATAGATACGAGGAAGATAGGAACCGTGAGCGGTGTAAAAG	
T	1051

FIG. 42

MEKVFLEKLQKTLHIPGGLLFYKGESGKTKTAFEFAGILCKENVPWGC	
GSCPSCKHVNELEEAFKGEIEDFKVYKDKDGKKHFVYLMGEHPDFVVI	100
PSGHYIKIEQIREVKNFAYVKPALSRKVI IIDDAHAMTSQAANALLKVL	
EEPPADTTFILTTNRRSAILPTILSRTFQVEFKGFSVKEVMEIAKVDEEI	200
AKLSGGSILKRAILLKENKDILNKVKEFLENEPLKVYKLASEFEKWEPEKQ	
KLFLEIMEELVSQKLTEEKDNYTYLLDTRLFKDGLARGVNEPLWLFTL	300
AVQAD	

FIG. 43

ATGAACTTCCTGAAAAAGTTCCTTTTACTGAGAAAAGCTCAAAAGTCTCC  
 TTACTTCGAAGAGTTCTACGAAGAAATCGATTTGAACCAGAAGGTGAAAG 100  
 ATGCAAGGTTTGTAGTTTTTGGACTGCGAAGCCACAGAACTCGACGTAAAG  
 AAGGCAAACTCCTTTCAATAGGTGCGGTTGAGGTTAAAAACCTGGAAAT 200  
 AGACCTCTCTAAATCTTTTACGAGATACTCAAAAGTGACGAGATAAAGG  
 CGGCGGAGATACATGGAATAACCAGGGAAGACGTTGAAAAGTACGGAAAG 300  
 GAACCAAAGGAAGTAATATACGACTTTCTGAAGTACATAAAGGGAAGCGT  
 TCTCGTTGGCTACTACGTGAAGTTTGACGTCTCACTCGTTGAGAAGTACT 400  
 CCATAAAGTACTTCCAGTATCCAATCATCAACTACAAGTTAGACCTGTTT  
 AGTTTCGTGAAGAGAGAGTACCAGAGTGGCAGGAGTCTTGACGACCTTAT 500  
 GAAGGAACTCGGTGTAGAAATAAGGGCAAGGCACAACGCCCTTGAAGATG  
 CCTACATAACCGCTCTTCTTTTCTAAAGTACGTTTACCCGAACAGGGAG 600  
 TACAGACTAAAGGATCTCCCGATTTTCCTT

FIG. 44

MNFLKKFLLLRKAQKSPYFEEFYEEIDLNQKVKDARFVVFDCEATELDVK  
 KAKLLSIGAVEVKNLEIDLKSFYEILKSDEIKAAEIHGITREDVEKYGK 100  
 EPKEVIYDFLKYIKGSVLVGYYVKFDVSLVEKYSIKYFQYPIINYLKDLF  
 SFVKREYQSGRSLDDLMKELGVEIRARHNALEDAYITALLFLKYVYPNRE 200  
 YRLKDLPIFL

FIG. 45

ATGCTCAATAAGGTTTTTATAATAGGAAGACTTACGGGTGACCCCGTTAT	
AACTTATCTACCGAGCGGAACGCCCGTAGTAGAGTTTACTCTGGCTTACA	100
ACAGAAGGTATAAAAACCAGAACGGTGAATTCAGGAGGAAAGTCACTTC	
TTTGACGTAAAGGCGTACGGAAAAATGGCTGAAGACTGGGCTACACGCTT	200
CTCGAAAGGATACCTCGTACTCGTAGAGGGAAGACTCTCCCAGGAAAAGT	
GGGAGAAAGAAGGAAAGAAGTTCTCAAAGGTCAGGATAATAGCGGAAAAC	300
GTAAGATTAATAAACAGGCCGAAAGGTGCTGAACCTTCAAGCAGAAGAAGA	
GGAGGAAGTTCCTCCCATTTGAGGAGGAAATTGAAAACTCGGTAAAGAGG	400
AAGAGAAGCCTTTTACCGATGAAGAGGACGAAATACCTTTTTAATTTTGA	
GGAGGTTAAAGTATGGTAGTGAGAGCTCCTAAGAAGAAAGTTTGTATGTA	500
CTGTGAACAAAAGAGAGAGCCAGATT	

**FIG. 46**

MLNKVFIIGRLTGDPVITYLPSGTPVVEFTLAYNRRYKNQNGEFQEESHF	
FDVKAYGKMAEDWATRFSGYLVLEGRLSQEKWEKEGKKFSKVRIIAEN	100
VRLINRPKGAELOAEIEEEVPPIEEEIEKLGKEEEKPFTDEEDEIPF	

**FIG. 47**

ATGCAATTTGTGGATAAACTTCCCTGTGACGAATCCGCCGAGAGGGCGGT	
TCTTGGCAGTATGCTTGAAGACCCCGAAAACATACCTCTGGTACTTGAAT	100
ACCTTAAAGAAGAAGACTTCTGCATAGACGAGCACAAGCTACTTTTCAGG	
GTTCTTACAAACCTCTGGTCCGAGTACGGCAATAAGCTCGATTTTCGTATT	200
AATAAAGGATCACCTTGAAAAGAAAACTTACTCCAGAAAATACCTATAG	
ACTGGCTCGAAGAACTCTACGAGGAGGCGGTATCCCTGACACGCTTGAG	300
GAAGTCTGCAAAATAGTAAAACAACGTTCCGCACAGAGGGCGATAATTCA	
ACTCGGTATAGAACTCATTACAAAGGAAAGGAAAACAAAGACTTTCACA	400
CATTAATCGAGGAAGCCCAGAGCAGGATATTTTCCATAGCGGAAAGTGCT	
ACATCTACGCAGTTTTTACCATGTGAAAGACGTTGCGGAAGAAGTTATAGA	500
ACTCATTTATAAATTCAAAGCTCTGACAGGCTAGTCACGGGACTCCCAA	
GCGGTTTTACGGAACTCGATCTAAAGACGACGGGATTCCACCCTGGAGAC	600
TTAATAATACTCGCCGCAAGACCCGGTATGGGGAAAACCGCCTTTATGCT	
CTCCATAATCTACAATCTCGCAAAAGACGAGGGAAAACCTCAGCTGTAT	700
TTTCCCTTGAAATGAGCAAGGAACAGCTCGTTATGAGACTCCTCTCTATG	
ATGTCGGAGGTCCCACTTTTCAAGATAAGGTCTGGAAGTATATCGAATGA	800
AGATTTAAAGAAGCTTGAAGCAAGCGCAATAGAACTCGCAAAGTACGACA	
TATACCTCGACGACACACCCGCTCTCACTACAACGGATTTAAGGATAAGG	900
GCAAGAAAGCTCAGAAAGGAAAAGGAAGTTGAGTTCGTGGCGGTGGACTA	
CTTGCAACTTCTGAGACCGCCAGTCCGAAAGAGTTCAAGACAGGAGGAAG	1000
TGGCAGAGGTTTCAAGAACTTAAAAGCCCTTGCAAAGGAACTTCACATT	
CCCGTTATGGCACTTGCGCAGCTCTCCCGTGAGGTGGAAAAGAGGAGTGA	1100
TAAAAGACCCAGCTTGCGGACCTCAGAGAATCCGGACAGATAGAACAGG	
ACGCAGACCTAATCCTTTTCTCCACAGACCCGAGTACTACAAGAAAAAG	1200
CCAAATCCCGAAGAGCAGGGTATAGCGGAAGTGATAATAGCCAAGCAAAG	
GCAAGGACCCACGGACATTGTGAAGCTCGCATTTATTAAGGAGTACACTA	1300
AGTTTGCAAACCTAGAAGCCCTTCTGAACAACCTCCTGAAGAAGAGGAA	
CTTTCGGAAATTATTGAAACACAGGAGGATGAAGGATTGGAAGATATTGA	1400
CTTCTGAAAATTAAGSTTTTATAATTTTATCTTGGCTATCCGGGGTAGCT	
CAATCGGCAGAGCGGGTGGCTG	1472

FIG. 48

MQFVDKLPCEESAERAVLGSMLLEDPENIPLVLEYLKEEDFCIDEHKLLFR	
VLTNLWSEYGNKLDVFLIKDHLEKKNLLQKIPIDWLEELYEEAVSPDTLE	100
EVCKIVKQRSAQRAIIQLGITSTQFYHVKDVAEEVIELIYKFKSSDRLVT	
GLPSGFTELDLKTTFHFGDLIIILARPFGMGKTAFLMSIIYNLAKDEGKP	200
SAVFSLEMSKEQLVMRLLSMMSEVPLFKIRSGSISNEDLKKLEASAIELA	
KYDIYLDLDDTPALTTDLRIRARKLRKEKEVEFVAVDYLQLLRPPVRKSSR	300
QEEVAEVSRLKALAKELHIPVMALAQLSREVEKRS DKRPQLADLRESGQ	
IEQDADLILFLHRPEYYKKKPNPEEQGIAEVIIAKQRQGP TDIVKLAFIK	400
EYTKFANLEALPEQPPEEEELSEIIETQEDEGFEDIDF	

FIG. 49

ATGTCCTCGGACATAGACGAACTTAGACGGGAAATAGATATAGTAGACGT	
CATTTCCGAATACTTAAACTTAGAGAAGGTAGGTTCCAATTACAGAACGA	100
ACTGTCCCTTTTACCCTGACGATACACCCTCCTTTTACGTGTCTCCAAGT	
AAACAAATATTCAAGTGTTTCGGTTGCGGGGTAGGGGAGACGCGATAAA	200
GTTTCGTTTCCCTTTACGAGGACATCTCCTATTTTGAAGCCGCCCTTGAAC	
TCGCAAAACGCTACGGAAAGAAATTAGACCTTGAAAAGATATCAAAGAC	300
GAAAAGGTATACGTGGCTCTTGACAGGGTTTGTGATTTCTACAGGGAAAG	
CCTTCTCAAAAACAGAGAGGCAAGTGAGTACGTAAAGAGTAGGGGAATAG	400
ACCCTAAAGTAGCGAGGAAGTTTGATCTTGGGTACGCACCTTCCAGTGAA	
GCACTCGTAAAGTCTTAAAAGAGAACGATCTTTTAGAGGCTTACCTTGA	500
AACTAAAAACCTCCTTTCTCCTACGAAGGGTGTTTACAGGGATCTCTTTC	
TTCCGGCGTGTCGTGATCCCGATAAAGGATCCGAGGGGAAGAGTTATAGGT	600
TTCCGGTGGAAGGAGGATAGTAGAGGACAAATCTCCAAGTACATAAACTC	
TCCAGACAGCAGGGTATTTAAAAAGGGGGAGAACTTATTCGGTCTTTACG	700
AGGCAAAGGAGTATATAAAGGAAGAAGGATTTGCGATACTTGTGGAAGGG	
TACTTTGACCTTTTGAGACTTTTTTCCGAGGGGAATAAGGAACGTTGTTGC	800
ACCCCTCGGTACAGCCCTGACCCAAAATCAGGCAAACCTCCTTTCCAAGT	
TCACAAAAAAGGTCTACATCCTTTACGACGGAGATGATGCGGGAAGAAAG	900
GCTATGAAAAGTGCCATTCCCCTACTCCTCAGTGACAGGAGTGGAAGTTTA	
TCCCGTTTACCTCCCCGAAGGATACGATCCCGACGAGTTTATAAAGGAAT	1000
TCGGGAAAGAGGAATTAAGAAGACTGATAAACAGCTCAGGGGAGCTCTTT	
GAAACGCTCATAAAAACCGCAAGGGAAAACCTTAGAGGAGAAAACGCGTGA	1100
GTTTCAGGTATTATCTGGGCTTTATTTCCGATGGAGTAAGGCGCTTTGCTC	
TGGCTTCGGAGTTTCACACCAAGTACAAAGTTCCTATGGAAATTTTATTA	1200
ATGAAAATTGAAAAAAATTCTCAAGAAAAAGAAATTAACTCTCCTTTAA	
GGAAAAATCTTCCTGAAAGGACTGATAGAATTAAAACCAAAAATAGACC	1300
TTGAAGTCCTGAACTTAAGTCCTGAGTTAAAGGAACTCGCAGTTAACGCC	
TTAAACGGAGAGGAGCATTACTTCCAAAAGAAGTTCTCGAGTACCAGGT	1400
GGATAACTTGAGAGAACTTTTTTAACAACATCCTTAGGGATTTACAAAAAT	
CTGGGAAAAAGAGGAAGAAAAGAGGGTTGAAAAATGTAAATACTTAATTA	1500
ACTTTAATAAATTTTTAGAGTTAGGA	

FIG. 50

MSSDIDELRREIDIVDVISEYLNLEKVGSNYRTNCPFHPDDTPSFYVSPS	
KQIFKFCGCGVGGDAIKFVSLYEDISYFEAALELAKRYGKKLDLEKISKD	100
EKVYVALDRVCDFYRESLLKNREASEYVKSREGIDPKVARKFDLGYAPSSE	
ALVKVLKENDLLEAYLETKNLLSPTKGVYRDLFLRRVVIPIKDPGRGVIG	200
FGGRRIVEDKSPKYINSPDSRVFKKGENLFLGLYEKEYIKEEGFAILVEG	
YFDLLRLFSEGIRNVVAPLGTALTQONQANLLSKFTKKVYILYDGDAGRK	300
AMKSAIPLLLSAGVEVYPVYLPEGYDPDEFIKEFGKEELRRLINSSGELF	
ETLIKTARENLEEKTRFRYYLGFISDGVRRFALASEFHTKYKVPMEILL	400
MKIEKNSQEKEIKLSFKEKIFLKGLIELKPKIDLEVLNLSPELKE LAVNA	
LNGBEHLKPKEVLEYQVDNLEKLFNNILRDLQKSGKKRKRGLKNVNT	498

FIG. 51

ATGCAAGATACCGCTACCTGCAGTATTTGTCAGGGGACGGGATTTCGTAAA	
GACCGAAGACAACAAGGTAAGGCTCTGCGAATGCAGGTTCAAGAAAAGGG	100
ATGTAAACAGGGAACTAAACATCCCAAAGAGGTACTGGAACGCCAACTTA	
GACACTTACCACCCCAAGAACGTATCCCAGAACAGGGCACTTTTGACGAT	200
AAGGGTCTTCGTCCACAACCTTCAATCCCGAGGAAGGGAAAGGGCTTACCT	
TTGTAGGATCTCCTGGAGTCGGCAAACTCACCTTGCGGTTGCAACATTA	300
AAAGCGATTTATGAGAAGAAGGGAATCAGAGGATACTTCTTCGATACGAA	
GGATCTAATATTCAGGTTAAAACACTTAATGGACGAGGGAAAGGATACAA	400
AGTTTTTAAAACTGTCTTAAACTCACCGGTTTTTGGTTCTCGACGACCTC	
GGTTCTGAGAGGCTCAGTGACTGGCAGAGGGAATCATCTCTTACATAAT	500
CACTTACAGGTATAACAACCTTAAGAGCACGATAATAACCACGAATTACT	
CACTCCAGAGGGAAGAAGAGAGTAGCGTGAGGATAAGTGCGGATCTTGCA	600
AGCAGACTCGGAGAAAACGTAGTTTCAAAAATTTACGAGATGAACGAGTT	
GCTCGTTATAAAGGGTTCCGACCTCAGGAAGTCTAAAAAGCTATCAACCC	700
CATCT	

FIG. 52

MQDTATCSICQGTGFVKTEDNKNVRLCECRFKKRDVNRELNIPKRYWNANL	
DTYHPKNVSQNRALLTIRVFVHNFNP EEGKGLTFVGSPGVGKTHLAVATL	100
KAIYEKKGIRGYFFDTKDLIFRLKHLMD EGKDTKFLKTVLNSPVLVDDL	
GSERLSDWQRELISYIIITYRYNNLKSTIIITNYSLQREEESSVRISADLA	200
SRLGENVVSKIYEMNELLVIKGSDLRKSKKLSTPS	

FIG. 53

ATGAAAAAGATTGAAAATTTGAAGTGAAAAATGTCTCGTTTAAAAGCCT	100
GGAAATAGATCCCGATGCAGGTGTGGTTCTCGTTTCCGTGGAAAAATTCT	
CCGAAGAGATAGAAGACCTTGTGCGTTTACTGGAGAAGAAGACGCGGTTT	200
CGAGTCATCGTGAACGGTGTTCAAAAAGTAACGGGGATCTAAGGGGAAA	
GATACTTTCCTTCTCAACGGTAATGTGCCTTACATAAAAGATGTTGTTT	300
TCGAAGGAAACAGGCTGATTCTGAAAGTGCTTGGAGATTTGCGCGGGAC	
AGGATCGCCTCCAACTCAGAAGCACGAAAAACAGCTCGATGAACTGCT	400
GCCTCCCGGAACAGAGATCATGCTGGAGGTTGTGGAGCCTCCGGAAGATC	
TTTTGAAAAAGGAAGTACCACAACCAGAAAAGAGAGAAGAACCAAAGGGT	500
GAAGAATTGAAGATCGAGGATGAAAACCACATCTTTGGACAGAAACCAG	
AAAGATCGTCTTCACCCCTCAAAAATCTTTGAGTACAACAAAAGACAT	600
CGGTGAAGGGCAAGATCTTCAAATAGAGAAGATCGAGGGGAAAAGAACG	
GTCTTCTGATTTACCTGACAGACGGAGAAGATTCTCTGATCTGCAAAGT	700
CTTCAACGACGTTGAAAAGGTGCAAGGGAAAGTATCGGTGGGAGACGTGA	
TCGTTGCCACAGGAGACCTCCTTCTCGAAAACGGGGAGCCACCTTTAC	800
GTGAAGGGAATCACAAAACCTCCCGAAGCGAAAAGGATGGACAAATCTCC	
GGTTAAGAGGGTGGAGCTCCACGCCCATACCAAGTTCAGCGATCAGGACG	900
CAATAACAGATGTGAACGAATATGTGAAACGAGCCAAGGAATGGGGCTTT	
CCCGCGATAGCCCTCACGGATCATGGGAACGTTCAGGCCATACCTTACTT	1000
CTACGACGCGGCGAAAGAAGCTGGAATAAAGCCCATTTCGGTATCGAAG	
CGTATCTGGTGAGTGACGTGGAGCCCGTCATAAGGAATCTCTCCGACGAT	1100
TCGACGTTTGGAGATGCCACGTTTCGTCTCTCGACTTCGAGACGACGGG	
TCTCGACCCGCAGGTGGATGAGATCATCGAGATAGGAGCGGTGAAGATAC	1200
AGGGTGGCCAGATAGTGGACGAGTACCACACTCTCATAAAGCCTTCCAGG	
GAGATCTCAAGAAAAAGTTTCGGAGATCACCGGAATCACTCAAGAGATGCT	1300
GGAAAACAAGAGAAGCATCGAGGAAGTTCTGCCGGAGTTTCTCGGTTTTTC	
TGGAAGATTCCATCATCGTAGCACACAACGCCAACTTCGACTACAGATTT	1400
CTGAGGCTGTGGATCAAAAAGTGATGGGATTGGACTGGGAAAGACCCTA	
CATAGATACGCTCGCCCTCGCAAAGTCCCTTCTCAAACCTGAGAAGCTACT	1500
CTCTGGATTCCGTTGTGGAAAAGCTCGGATTGGGTCCCTTCGGCACACC	
AGGGCCCTGGATGACGCGAGGGTCACCGCTCAGGTTTTCTCAGGTTTCGT	1600
TGAGATGATGAAGAAGATCGGTATCACGAAGCTTTCAGAAATGGAGAAGT	
TGAAGGATACGATAGACTACACCGCGTTGAAACCCTTCCACTGCACGATC	1700
CTCGTTCAGAACAAAAAGGGATTGAAAAACCTATACAAACTGGTTTTCTGA	
TTCTATATAAAGTACTTCTACGGTGTTCGGAGGATCCTCAAAGTGAGC	1800
TCATCGAGAACAGAGAAGGACTGCTCGTGGGTAGCGCGTGTATCTCCGGT	
GAGCTCGGACGTGCCGCCCTCGAAGGAGCGAGTGATTGAGAACTCGAAGA	1900
GATCGCGAAGTTCTACGACTACATAGAAGTCATGCCGCTCGACGTTATAG	
CCGAAGATGAAGAAGACCTAGACAGAGAAAGACTGAAAGAAGTGTAACGA	2000
AAACTCTACAGAATAGCGAAAAAATTGAACAAGTTCGTCTCATGACCGG	
TGATGTTTCAATTTCTCGATCCCGAAGATGCCAGGGGCAGAGCTGCACTTC	2100
TGGCACCTCAGGGAAACAGAACTTCGAGAATCAGCCCGCACTCTACCTC	
AGAACGACCGAAGAAATGCTCGAGAAGGCGATAGAGATATTGGAAGATGA	2200
AGAGATCGCGAGGGAAGTCGTGATAGAGAATCCCAACAGAATAGCCGATA	
TGATCGAGGAAGTGACCGCTCGAGAAAAAATTCACCCGCCGATCATA	2300
GAGAACGCCGATGAAATAGTGAGAAACCTCACCATGAAGCGGGCGTACGA	
GATCTACGGTGATCCGCTTCCCGAAATCGTCCAGAAGCGTGTGGAAAAGG	

FIG. 54A



AACTGAACGCCATCATAAATCATGGATACGCCGTTCTCTATCTCATCGCT	2400
CAGGAGCTCGTTCAGAAATCTATGAGCGATGGTTACGTGGTTGGATCCAG	
AGGATCCGTCGGGTCTTCACTCGTGGCCAATCTCCTCGGAATAACAGAGG	2500
TGAATCCCCTACCACCACATTACAGGTGTCCAGAGTGCAAATACTTTGAA	
GTTGTCGAAGACGACAGATACGGAGCGGGTTACGACCTTCCCAACAAGAA	2600
CTGTCCAAGATGTGGGGCTCCTCTCAGAAAAGACGGCCACGGCATAACCGT	
TTGAAACGTTTCATGGGGTTCGAGGGTGACAAGGTCCCCGACATAGATCTC	2700
AACTTCTCAGGAGAGTATCAGGAACGTGCTCATCGTTTTGTGGAAGAACT	
CTTCGGTAAAGACCACGTCTATAGGGCGGGAACCATAAACACCATCGCGG	2800
AAAGAAGTGCGGTGGGTACGTGAGAAGCTACGAAGAGAAAACCGGAAAG	
AAGCTCAGAAAGGCGGAAATGGAAAGACTCGTTTCCATGATCACGGGAGT	2900
GAAGAGAACGACGGGTGAGCACCAGGGGGGCTCATGATCATAACCGAAAG	
ACAAAGAAGTCTACGATTTCACTCCCATAACAGTATCCAGCCAACGATAGA	3000
AACGCAGGTGTGTTCAACCACGCACTTCGCATACGAGACGATCCATGATGA	
CCTGGTGAAGATAGATGCGCTCGGCCACGATGATCCCACTTTTCATCAAGA	3100
TGCTCAAGGACCTCACCGGAATCGATCCCATGACGATTCCCATGGATGAC	
CCCGATACGCTCGCCATATTCAAGTTCTGTGAAGCCTCTTGGTGTGGATCC	3200
CGTTGAGCTGGAAAGCGATGTGGGAACGTACGGAATTCCGGAGTTCGGAA	
CCGAGTTTGTGAGGGGAATGCTCGTTGAAACGAGACCAAAGAGTTTCGCC	3300
GAGCTTGTGAGAATCTCAGGACTGTACACGGTACGGACGTCTGGTTGAA	
CAACGCACGTGATTGGATAAACCTCGGCTACGCCAAGCTCTCCGAGGTTA	3400
TCTCGTGTAGGGACGACATCATGAACTTCCTCATAACAAAGGAATGGAA	
CCGTCACTTGCCTTCAAGATCATGGAAAACGTCAGGAAGGGAAAGGGTAT	3500
CACAGAAGAGATGGAGAGCGAGATGAGAAGGCTGAAGGTTCCAGAATGGT	
TCATCGAATCCTGTAAAAGGATCAAATATCTCTTCCCGAAAGCTCACGCT	3600
GTGGCTTACGTGAGTATGGCCTTCAGAATTGCTTACTTCAAGGTTCACTA	
TCCTCTTCAGTTTTACGCGGCGTACTTCACGATAAAAGGTGATCAGTTTCG	3700
ATCCGGTTCTCGTACTCAGGGGAAAAGAAGCCATAAAGAGGCGCTTGAGA	
GAACTCAAAGCGATGCCTGCCAAAGACGCCCAGAAGAAAAACGAAGTGAG	3800
TGTTCTGGAGGTTGCCCTGGAAATGATACTGAGAGGTTTTTCTTCTTCTAC	
CGCCCGACATCTTCAAATCCGACGCGAAGAAATTTCTGATAGAAGGAAAC	3900
TCGCTGAGAATTCCGTTCAACAAACTTCCAGGACTGGGTGACAGCGTTGC	
CGAGTCGATAATCAGAGCCAGGGAAGAAAAGCCGTTCACTTCGGTGGGAAG	4000
ATCTCATGAAGAGGACCAAGGTCAACAAAAATCACATAGAGCTGATGAAA	
AGCCTGGGTGTTCTCGGGGACCTTCAGAGACGGAACAGTTCACGCTTTT	4100

C

FIG. 54B

MKKIENLKWKNVSFKSLEIDPDAGVVLVSVEKFSEEIEDLVRLLEKKTRF	
RVIVNGVQKSNGLRGKILSLLNGNVPYIKDVVFEGNRLILKVLGDFARD	100
RIASKLRSTKKQLDELLPPGTEIMLEVVEPPEDLLKKEVPQPEKREEPKG	
EELKIEDENHIFGQKPRKIVFTPSKIFEYNKKTSVKGKIFKIEKIEGKRT	200
VLLIYLTGDGDSLICKVFNDVEKVEGKVSVDVIVATGDLLLENGEPTLY	
VKGITKLPEAKRMDKSPVKRVELHAHTKFSQDAITDVNEYVKRAKEWGF	300
PAIALTDHGNVQAIPIFYDAAKEAGIKPIFGIEAYLVSDVEPVIRNLSDD	
STFGDATFVVLDFETTGLDPQVDEIIIEIGAVKIQGGQIVDEYHTLIKPSR	400
EISRKSSSEITGITQEMLENKRSIEEVLPEFLGFLEDSEIVAHNANFDYRF	
LRLWIKKVMGLDWERPYIDTLALAKSLLKLSYSLSVVEKLGGLGPFRRH	500
RALDDARVTAQVFLRFVEMMKIGITKLSEMEKLDITIDYTALKPFHCTI	
LVQNKGLKNLYKLVSDSYIKYFYGVPRILKSELIENTREGLLVGSACISG	600
ELGRAALEGASDSELEEIAKFYDYIEVMPLDVIAEDEEDLDRERLKEVYR	
KLYRIAKKLNKFVVMGTGDVHFLDPEDARGRAALLAPQGNRNFENQPALYL	700
RTTEEMLEKAIEIFEDEEIIAREVVIENTPNRIADMIEEVQPLEKKLHPPII	
ENADEIVRNLTMKRAYEIIYGDPLPEIVQKRVEKELNAIINHGYAVLYLIA	800
QELVQKSMDSGYVVGSRGSSLVANLLGITEVNPLPPHYRCPECKYFE	
VVEDDRYGAGYDLPNKNCPRCGAPLRKDGHGIPFETFMGFEGDKVPDIDL	900
NFSGEYQERAHRFVEELFGKDHVYRAGTINTIAERSAVGYVRSYEEKTGK	
KLRKAEMERLVSMITGVKRTTGQHPGGLMIIPKDKEVYDFTPIQYPANDR	1000
NAGVFTTHFAYETIHDDLVKIDALGHDDPTFIKMLKDLTGIDPMTIPMDD	
PDTLAI FSSVKPLGVD PVELES DVGTYGIPEFGTEFVRGMLVETRPKSFA	1100
ELVRISGLSHGTDVWLNNARDWINLGYAKLSEVISCRDDIMNFLIHKGME	
PSLAFKIMENVRKKGKITEEMESEMRRLKVPEWFIESCKRIKYLFPKAHA	1200
VAYVSMAFRIAYFKVHYPLQFYAAYFTIKGDQFDPVLVLRGKEAIKRRLR	
ELKAMPAKDAQKKNEVSVLEVALEMILRGFSFLPPDIFKSDAKKFLIEGN	1300
SLRIPFNKLPGLGDSVAESIIRAREEKPFSTSVEDLMKRTKVNKNHIELMK	
SLGVLGDLPETEQFTLF	1367

FIG. 55

GTGCTCGCCATGATATGGAACGACACCGTTTTTTGCGTCGTAGACACAGA	
AACCACGGGAACCGATCCCTTTGCCGGAGACCGGATAGTTGAAATAGCCG	100
CTGTTCTGTCTTCAAGGGGAAGATCTACAGAAACAAAGCGTTTCACTCT	
CTCGTGAATCCCAGAATAAGAATCCCTGCGCTGATTACAGAAAGTTCACGG	200
TATCAGCAACATGGACATCGTGGAAGCGCCAGACATGGACACAGTTTACG	
ATCTTTTCAGGGATTACGTGAAGGGAACGGTGCTCGTGTTTCACAACGCC	300
AACTTCGACCTCACTTTTCTGGATATGATGGCAAAGGAAACGGGAAACTT	
TCCAATAACGAATCCCTACATCGACACACTCGATCTTTCAGAAGAGATCT	400
TTGGAAGGCCTCATTCTCTCAAATGGCTCTCCGAAAGACTTGGAATAAAA	
ACCACGATACGGCACCGTGCTCTTCCAGATGCCCTGGTGACCGCAAGAGT	500
TTTTGTGAAGCTTGTTGAATTTCTTGGTGAAAACAGGGTCAACGAATTCA	
TACGTGGAAAACGGGGG	567

**FIG. 56**

MLAMIWNDTVFCVVDTETTGTDPFAGDRIVEIAAVPVFKGKIYRNKAFHS	
LVNPRIRIPALIQKVHGISNMDIVEAPDMDTVYDLFRDYVKGTVLVFHNA	100
NFDLTFLDMMAKETGNFPITNPYIDTLDLSEEIFGRPHSLKWLSEIRLGK	
TTIRHRALPDALVTARVFVKLVEFLGENRVNEFIRGKRG	189

**FIG. 57**

GTGGAAGTTCTTTACAGGAAGTACAGGCCAAAGACTTTTTCTGAGGTTGT	
CAATCAGGATCATGTGAAGAAGGCAATAATCGGTGCTATTCAGAAGAACA	100
GCGTGGCCACGGATACATATTCGCCGGTCCGAGGGGAACGGGGAAGACT	
ACTCTTGCCAGAATTCTCGCAAATCCCTGAACTGTGAGAACAGAAAGGG	200
AGTTGAACCCTGCAATTCCTGCAGAGCCTGCAGAGAGATAGACGAGGGAA	
CCTTCATGGACGTGATAGAGCTCGACGCGGCCTCCAACAGAGGAATAGAC	300
GAGATCAGAAGAATCAGAGACGCCGTTGGATACAGGCCGATGGAAGGTAA	
ATACAAAGTCTACATAATAGACGAAGTTCACATGCTCACGAAAGAAGCCT	400
TCAACGCGCTCCTCAAAACACTCGAAGAACCTCCTTCCCACGTCGTGTTC	
GTGCTGGCAACGACAAACCTTGAGAAGGTTCTTCCCACGATTATCTCGAG	500
ATGTCAGGTTTTTCGAGTTCAGAAACATTCCCGACGAGCTCATCGAAAAGA	
GGCTCCAGGAAGTTGCGGAGGCTGAAGGAATAGAGATAGACAGGGAAGCT	600
CTGAGCTTCATCGCAAAAAGAGCCTCTGGAGGCTTGAGAGACGCGCTCAC	
CATGCTCGAGCAGGTGTGGAAGTTCCTCGGAAGGAAAGATAGATCTCGAGA	700
CGGTACACAGGGCGCTCGGGTTGATACCGATACAGGTTGTTTCGCGATTAC	
GTGAACGCTATCTTTTCTGGTGATGTGAAAAGGGTCTTCACCGTTCTCGA	800
CGACGTCTATTACAGCGGGAAGGACTACGAGGTGCTCATTACAGGAAGCAG	
TCGAGGATCTGGTCGAAGACCTGGAAAGGGAGAGAGGGGTTTACCAGGTT	900
TCAGCGAACGATATAGTTCAGGTTTCGAGACAACTTCTGAATCTTCTGAG	
AGAGATAAAGTTCGCCGAAGAAAAACGACTCGTCTGTAAAGTGGGTTTCGG	1000
CTTACATAGCGACGAGGTTCTCCACCACAAACGTTACAGGAAAACGATGTC	
AGAGAAAAAAACGATAATTCAAATGTACAGCAGAAAGAAGAGAAGAAAGA	1100
AACGGTGAAGGCAAAGAAGAAAAACAGGAAGACAGCGAGTTCGAGAAAC	
GCTTCAAAGAACTCATGGAAGAACTGAAAGAAAAGGGCGATCTCTCTATC	1200
TTTGTGCTCTCAGCCTCTCAGAGGTGCAGTTTGACGGAGAAAAGGTGAT	
TATTTCTTTTGATTCATCGAAAGCTATGCATTACGAGTTGATGAAGAAAA	1300
AACTGCCTGAGCTGGAAAACATTTTTTCTAGAAAACTCGGGAAAAAAGTA	
GAAGTTGAACTTCGACTGATGGGAAAAGAAGAAACAATCGAGAAGGTTTC	1400
TCAGAAGATCCTGAGATTGTTTGAACAGGAGGGA	

FIG. 58

MEVLYRKYPKTFSEVVNQDHVKKAIIGAIQKNSVAHG YIFAGPRGTGKT	
TLARILAKSLNCENRKGVEPCNSCRACREIDEGTFMDVIELDAASNRGID	100
EIRRIRDAVG YRPMEGKYKVYIIDEVHMLTKEAFNALLKTLEPPSHVVF	
VLATTNLEKVPPTIISRCQVFEFRNIPDELIEKRLQEVAAEAGIEIDREA	200
LSFIAKRASGGLRDALTMLEQVWKFSEGKIDLETVHRALGLIPIQVVRDY	
VNAIFSGDVKRVFTVLDDVYYS GDKDYEVLIQEAVEDLVEDLERERGVYQV	300
SANDIVQVSRQLLNLLREIKFAEEKRLVCKVGSAYIATRFSTTNVQENDV	
REKNDNSNVQOKEEKKETVKAKEEKQEDSEFEKRFKELMEELKEKGDL SI	400
FVALSLSEVQFDGEKVIISFDSSKAMHYELMKKKLPELENIFSRKLGKKV	
EVELRLMGKEETIEKVSQKILRLFEQEG	478

FIG. 59

ATGAAAGTAACCGTCACGACTCTTGAATTGAAAGACAAAATAACCATCGC	100
CTCAAAGCGCTCGCAAAGAAATCCGTGAAACCCATTCTTGCTGGATTTC	
TTTTCGAAGTGAAAGATGGAAATTTCTACATCTGCGCGACCGATCTCGAG	200
ACCGGAGTCAAAGCAACCGTGAATGCCGCTGAAATCTCCGGTGAGGCACG	
TTTTGTGGTACCAGGAGATGTCATTTCAGAAGATGGTCAAGGTTCTCCCAG	300
ATGAGATAACGGAACCTTTCTTTAGAGGGGGATGCTCTTGTTATAAGTTCT	
GGAAGCACCGTTTTTCAGGATCACCACCATGCCGCGGACGAATTTCCAGA	400
GATAACGCCTGCCGAGTCTGGAATAACCTTCGAAGTTGACACTTCGCTCC	
TCGAGGAAATGGTTGAAAAGGTCATCTTCGCCGCTGCCAAAGACGAGTTC	500
ATGCGAAATCTGAATGGAGTTTTCTGGGAACTCCACAAGAATCTTCTCAG	
GCTGGTTGCAAGTGATGGTTTTCAGACTTGCACTTGCTGAAGAGCAGATAG	600
AAAACGAGGAAGAGGCGAGTTTTCTTGCTCTCTTTGAAGAGCATGAAAGAA	
GTTCAAACCGTGCTGGACAACACAACGGAGCCGACTATAACGGTGAGGTA	700
CGATGGAAGAAGGGTTTTCTCTGTCGACAAATGATGTAGAAACGGTGATGA	
GAGTGGTCGACGCTGAATTTCCCGATTACAAAAGGGTGATCCCCGAACT	800
TTCAAACGAAAGTGGTGGTTTTCCAGAAAAGAACTCAGGGAATCTTTGAA	
GAGGGTGATGGTGATTGCCAGCAAGGGAAGCGAGTCCGTGAAGTTCGAAA	900
TAGAAGAAAACGTTATGAGACTTGTGAGCAAGAGCCCGATTATGGAGAA	
GTGGTCGATGAAGTTGAAGTTCAAAAAGAAGGGGAAGATCTCGTGATCGC	1000
TTTCAACCCGAAGTTCATCGAGGACGTTTTGAAGCACATTGAGACTGAAG	
AAATCGAAATGAACCTCGTTGATTCTACCAGTCCATGTGAGATAAATCCA	1098
CTCGATATTTCTGGATACCTTTACATAGTGATGCCCATCAGACTGGCA	

FIG. 60

MKVTVTTLLELKDKITIASKALAKKSVKPILAGFLFEVKDGNFYICATDLE	100
TGVKATVNAAEISGEARFVVPGDVIQKMVKVLPDEITELSLEGDALVISS	
GSTVFRITTMPADEFPEITPAESGITFEVDTSLLEEMVEKVIFAAKDEF	200
MRNLNGVFWELHKNLLRLVASDGFRLALAEQIENEEEASFLLSLKSMKE	
VQNVLDNTTEPTITVRYDGRRVSLSTNDVETVMRVVDAEFPDYKRVIPET	300
FKTKVVSRKELRESLKRVMVIAASKGSESVKFEIEENVMLVSKSPDYGE	
VVDEVEVQKEGEDLVIAFNPKFIEDVLKHIEETEEIEMNFVDSTSPCQINP	366
LDISGYLYIVMPIRLA	

FIG. 61

ATGCCAGTCACGTTTCTCACAGGTA	100
CTGCGAGAACTCAGAAAGGAAGAATT	
GATAAAGAACTCCTGAAGGATGGTAAC	
CGTGGAGTACATAAGGATCCATC	100
CGGAGGATCCCGACAAGATCGATTTC	
CATAAGGTCTTTACTCAGGACAAAG	
ACGATCTTTTCCAACAAGACGATCATT	200
GACATCGTCAATTTTCGATGAGTG	
GAAAGCACAGGAGCAGAAGCGTCTCG	
TGAACTTTTGAAAAACGTACCGG	
AAGACGTTTCATATCTTCATCCGTTCT	300
CAAAAAACAGGTGGAAAGGGAGTA	
GCGCTGGAGCTTCCGAAGCCATGGGAA	
ACGGACAAGTGGCTTGAGTGGAT	
AGAAAAGCGCTTCAGGGAGAATGGTTT	400
GTGCTCATCGATAAAGATGCCCTTC	
AGCTGTTTTTCTCCAAGGTTGGAACGA	
ACGACCTGATCATAGAAAGGGAG	
ATTGAAAAACTGAAAGCTTATTCCGAG	500
GACAGAAAGATAACGGTAGAAGA	
CGTGGAAGAGGTCGTTTTTACCTATCA	
GACTCCGGGATACGATGATTTTT	
GCTTTGCTGTTTCCGAAGGAAAAAGGA	600
AGCTCGCTCACTCTCTTCTGTGCG	
CAGCTGTGGAACCACAGAGTCCGTGGT	
GATTGCCACTGTCCTTGCGAA	
TCACTTCTTGATCTCTTCAAAATCCTC	700
GTGACAAAGAAAAGAT	
ACTACACCTGGCCTGATGTGTCCAGG	
GTGTCCAAAGAGCTGGGAATTCCC	
GTTCTCGTGTGGCTCGTTTCCTCGGT	800
TTCTCCTTTAAGACCTGGAAATT	
CAAGGTGATGAACCACCTCCTCTACTA	
CGATGTGAAGAAGGTTAGAAAGA	
TACTGAGGGATCTCTACGATCTGGACA	900
GAGCCGTGAAAAGCGAAGAAGAT	
CCAAAACCGTTCTTCCACGAGTTCATA	
GAAAGAGGTGGCACTGGATGTATA	
TTCTCTTCAGAGAGATGAAGAA	972

FIG. 62

MPVTFLTGTAEQKEELIKKLLKDG	
NVEYIRIHPEDPKIDFIRSLRLTK	
TIFSNKTIIDIVNFDEWKAQEQKRL	100
VELLKNVPEDVHIFIRSQKTGGKGV	
ALELPKPWETDKWLEWIEKRFRENG	
LLIDKDALQLFFSKVGTNDLI IERE	200
IEKLKAYSEDRKITVEDVEEVVFTY	
QTPGYDDFCFAVSEGKRKLAHSLLS	
QLWKTTESVVIATVLANHFLDLFKIL	300
VLVTKKRYTWPDVSRVSKELGIP	
VPRVARFLGFSFKTWKFKVMNHL	
LYYDVKKVRKILRDLYDLRAVKSEED	
PKPFFHEFIEEVALDVYSLQRDEE	

FIG. 63

ATGAACGATTTGATCAGAAAGTACGCTAAAGATCAACTGGAACTTTGAA	
AAGGATCATAGAAAAGTCTGAAGGAATATCCATCCTCATAAATGGAGAAG	100
ATCTCTCGTATCCGAGAGAAGTATCCCTTGAACCTCCCGAGTACGTGGAG	
AAATTTCCCCCGAAGGCCTCGGATGTTCTGGAGATAGATCCCGAGGGGGA	200
GAACATAGGCATAGACGACATCAGAACGATAAAGGACTTCCTGAACTACA	
GCCCCGAGCTCTACACGAGAAAGTACGTGATAGTCCACGACTGTGAAAGA	300
ATGACCCAGCAGGCGGCGAACGCGTTTCTGAAGGCCCTTGAAGAACCACC	
AGAATACGCTGTGATCGTTCTGAACACTCGCCGCTGGCATTATCTACTGC	400
CGACGATAAAGAGCCGAGTGTTTCTGAGTGGTTGTGAACGTTCCAAAGGAG	
TTCAGAGATCTCGTGAAAGAGAAAATAGGAGATCTCTGGGAGGAACTTCC	500
ACTTCTTGAGAGAGACTTCAAACGGCTCTCGAAGCCTACAAACTTGGTG	
CGGAAAAACTTTCTGGATTGATGGAAAGTCTCAAAGTTTGGAGACGGAA	600
AAACTCTTGAAAAAGGTCCTTTCAAAGGCCTCGAAGGTTATCTCGCATG	
TAGGGAGCTCCTGGAGAGATTTTCAAAGGTGGAATCGAAGGAATTCTTTG	700
CGCTTTTGGATCAGGTGACTAACACGATAACAGGAAAAGACGCGTTTCTT	
TTGATCCAGAGACTGACAAGAATCATTCTCCACGAAAACACATGGGAAAG	800
CGTTGAAGATCAAAAAGCGTGTCTTTCCTCGATTCAATTCTCAGGGTGA	
AGATAGCGAATCTGAACAACAACTCACTCTGATGAACATCCTCGCGATA	900
CACAGAGAGAGAAAGAGAGGTGTCAACGCTTGGAGC	

FIG. 64

MNDLIRKYAKDQLETLKRIIEKSEGISILINGEDLSYPREVSLELPEYVE	
KFPPKASDVLEIDPEGENIGIDDIRTIKDFLNYSPELYTRKYVIVHDCER	100
MTQQAANAFLKALEEPPEYAVIVLNTRRWYLLPTIKSRVFRVVVNPKE	
FRDLVKEKIGDLWEELPLLERDFKTALEAYKLGAEKLSGLMESLKVLETE	200
KLLKKVLSKLEGYLACRELLERFSKVESKEFFALFDQVTNTITGKDAFL	
LIQRLTRIILHENTWESVEDKSVSFLDSILRVKIANLNNKLTLMNILAIH	300
RERKRGVNAWS	

FIG. 65

ATGTCCTTTCTTCAACAAGATCATACTCATAGGAAGACTCGTGAGAGATCC	
CGAAGAGAGATACACGCTCAGCGGAACCTCCAGTCACCACCTTCACCATAG	100
CGGTGGACAGGGTTCCCAGAAAGAACGCGCCGGACGACGCTCAAACGACT	
GATTTCTTCAGGATCGTCACCTTTGGAAGACTGGCAGAGTTCGCTAGAAC	200
CTATCTCACCAAAGGAAGGCTCGTTCTCGTCGAAGGTGAAATGAGAATGA	
GAAGATGGGAAACACCCACTGGAGAAAAGAGGGTATCTCCGGAGGTTGTC	300
GCAAACGTTGTTAGATTATGGACAGAAAACCTGCTGAAACAGTTAGCGA	
GACTGAAGAGGAGCTGGAAATACCGGAAGAAGACTTTTCCAGCGATACCT	400
TCAGTGAAGATGAACCACCATT	

**FIG. 66**

MSFFNKIILIGRLVRDPEERYTTLSGTPVTTFTTIAVDRVPRKNAPDDAQT	
DFFRIVTFGRLAEFARTYLTGRLVLVEGEMRMRRWETPTGEKRVSPVV	100
ANVVRFMDRKPAETVSETEEELEIPEEDFSSDTFSEDEPPF	

**FIG. 67**



ATGCGTGTTCCCCCGCACAACTTAGAGGCCGAAGTTGCTGTGCTCGGAAG	
CATATTGATAGATCCGTCCGTAATAAACGACGTTCTTGAAATTTTGAGCC	100
ACGAAGATTTCTATCTGAAAAAACACCAACACATCTTCAGAGCGATGGAA	
GAGCTTTACGACGAAGGAAAACCGGTGGACGTGGTTTCCGTCTGTGACAA	200
GCTTCAAAGCATGGGAAAACCTCGAGGAAGTAGGTGGAGATCTGGAAGTGG	
CCCAGCTCGCTGAGGCTGTGCCCAGTTCTGCACACGCACTTCACTACGCG	300
GAGATCGTCAAGGAAAAATCCATTCTGAGGAACTCATTGAGATCTCCAG	
AAAAATCTCAGAAAGTGCCTACATGGAAGAAGATGTGGAGATCCTGCTCG	400
ACAACGCAGAAAAGATGATCTTCGAGATCTCAGAGATGAAAACGACAAAA	
TCCTACGATCATCTGAGAGGCATCATGCACCGGGTGTGAAAACCTGGA	500
GAACTTCAGGGAAAGAGCCAACCTTATAGAACCCGGTGTGCTCATAACGG	
GACTACCAACGGGATTCAAAAGTCTGGACAAACAGACCACAGGGTTCCAC	600
AGCTCCGATCTGGTGATAATAGCAGCGAGACCCTCCATGGGAAAAACCTC	
CTTCGCACTCTCAATAGCGAGGAACATGGCTGTCAATTTGAAATCCCCG	700
TCGGAATATTCACTCTCGAGATGTCCAAGGAACAGCTCGCTCAAAGACTA	
CTCAGCATGGAGTCCGGTGTGGATCTTTACAGCATCAGAACAGGATACCT	800
GGATCAGGAGAAGTGGGAAAGACTCACAATAGCGGCTTCTAAACTCTACA	
AAGCACCCATAGTTGTGGACGATGAGTCACTCCTCGATCCGCGATCGTTG	900
AGGGCAAAAGCGAGAAGGATGAAAAAGAATACGATGTAAAAGCCATTTT	
TGTCGACTATCTCCAGCTCATGCACCTGAAAGGAAGAAAAGAAAGCAGAC	1000
AGCAGGAGATATCCGAGATCTCGAGATCTCTGAAGCTCCTTGCGAGGGAA	
CTCGACATAGTGGTGATAGCGCTTTCACAGCTTTCGAGGGCCGTAGAACA	1100
GAGAGAAGACAAAAGACCGAGGCTGAGTGACCTCAGGGAATCCGGTGCGA	
TAGAACAGGACGCAGACACAGTCATCTTCATCTACAGGGAGGAATATTAC	1200
AGGAGCAAAAATCCAAAGAGGAAAGCAAGCTTCACGAACCTCACGAAGC	
TGAAATCATAATAGGTAAACAGAGAAACGGTCCCCTTGGAACGATCACTC	1300
TGATCTTCGACCCCAGAACGGTTACGTTCCATGAAGTCGATGTGGTGCAT	
TCA	1353

FIG. 68

MRVPPHNLEAEVAVLGSILIDPSVINDVLEILSHEDFYLLKKHQHIFRAME	
ELYDEGKPDVVSVCDKLQSMGKLEEVGGDLEVAQLAEAVPSSAHALHYA	100
EIVKEKSILRKLEISRKISESAYMEEDVEILLDNAEKMIFEISEMKTTK	
SYDHLRGIMHRVFENLENFRERANLIEPGVLITGLPTGFKSLDKQTTGFH	200
SSDLVIIAARPSMGKTSFALSIARNMAVNFEIPVGIFSLEMSKEQLAORL	
LSMESGVDLYSIRTGYLDQEKWERLTIAASKLYKAPIVVDDDESLLDPRSL	300
RAKARMKKEYDVKAIFVDYLQLMHLKGRKESRQOEISEISRSLKLLARE	
LDIVVIALSQLSRAVEQREDKRPRLSLDRESGAIEQDADTVIFIFYREEYY	400
RSKKSKEESKLHEPHEAEIIIGKQRNGPVGTTITLIFDPRTVTFHEVDVVH	
S	451

FIG. 69

GTGATTCTCGAGAGGTCATCGAGGAAATAAAAGAAAAGGTTGACATCGT	
AGAGGTCATTTCCGAGTACGTGAATCTTACCCGGGTAGGTTCTCTCTACA	100
GGGCTCTCTGTCCCTTTCATTTCAGAAACCAATCCTTCTTTCTACGTTTCAT	
CCGGGTTTGAAGATATACCATTTGTTTCGGCTGCGGTGCGAGTGGAGACGT	200
CATCAAATTTCTTCAAGAAATGGAAGGGATCAGTTTCCAGGAAGCGCTGG	
AAAGACTTGCCAAAAGAGCTGGGATTGATCTTTCTCTCTACAGAACAGAA	300
GGGACTTCTGAATACGGAAAATACATTCGTTTGTACGAAGAAACGTGGAA	
AAGGTACGTCAAAGAGCTGGAGAAATCGAAAGAGGCCAAAAGACTATTTAA	400
AAAGCAGAGGCTTCTCTGAAGAAGATATAGCAAAGTTTCGGCTTTGGGTAC	
GTCCCAAGAGATCCAGCATCTCTATAGAAGTTGCAGAAGGCATGAACAT	500
AACACTGGAAGAACTTGTTCAGATACGGTATCGCGCTGAAAAAGGGTGATC	
GATTCGTTGATAGATTCTGAAGGAAGAATCGTTGTTCCAATAAAGAACGAC	600
AGTGGTCATATTGTGGCTTTTGGTGGGCGTGCTCTCGGCAACGAAGAACC	
GAAGTATTTGAACTCTCCAGAGACCAGGTATTTTTTCGAAGAAGAAGACCC	700
TTTTTCTCTTCGATGAGGCGAAAAAGTGGCAAAAGAGGTTGGTTTTTTC	
GTCATCACCGAAGGCTACTTCGACGCGCTCGCATTTCAGAAAGGATGGAAT	800
ACCAACGGCGGTTCGCTGTTCTTGGGGCGAGTCTTTCAGAGAGGCGATTTC	
TAAAACCTTTCGGCGTATTTCGAAAAACGTCATACTGTGTTTCGATAATGAC	900
AAAGCAGGCTTCAGAGCCACTCTCAAATCCCTCGAGGATCTCCTAGACTA	
CGAATTCAACGTGCTTGTGGCAACCCCTCTCCTTACAAAGACCCAGATG	1000
AACTCTTTCAGAAAGAAGGAGAAGGTTTCATTGAAAAAGATGCTGAAAAAC	
TCGCGTTTCGTTTCGAATATTTTCTGGTGACGCTGGTGAGGTCTTCTTTGA	1100
CAGGAACAGCCCCGCGGGTGTGAGATCCTACCTTTCTTTCCTCAAAGGTT	
GGGTCCAAAAGATGAGAAGGAAAGGATATTTGAAACACATAGAAAATCTC	1200
GTGAATGAGGTTTCATCTTCTCTCCAGATACCAGAAAACCAGATTTTGAA	
CTTTTTTTGAAAGCGACAGGTCTAACTATGCCTGTTTCATGAGACCAAGT	1300
CGTCAAAGGTTTACGATGAGGGGAGAGGACTGGCTTATTTGTTTTTGAAC	
TACGAGGATTTGAGGGAAAAGATTCTGGAACCTGGACTTAGAGGTACTGGA	1400
AGATAAAAACGCGAGGGAGTTTTTCAAGAGAGTCTCACTGGGAGAAGATT	
TGAACAAAGTCATAGAAAACCTCCCAAAGAGCTGAAAGACTGGATTTTT	1500
GAGACAATAGAAAGCATTCTCTCCTCAAAGGATCCCGAGAAATCTCTCGG	
TGACCTCTCCGAAAAGTTGAAAATCCGACGGATAGAGAGACGTATCGCAG	1600
AAATAGATGATATGATAAAGAAAGCTTCAAACGATGAAGAAAGGCGTCTT	
CTTCTCTCTATGAAAGTGGATCTCCTCAGAAAAATAAAGAGGAGG	1695

FIG. 70

MIPREVIEEIKEKVDIVEVISEYVNLTRVGSSYRALCPFHSETNPSFYVH	
PGLKIYHCFGCGASGDVIKFLQEMEGISFQEALERLAKRAGIDLSLYRTE	100
GTSEYGKYIRLYEETWKRYVKELEKSKEAKDYLSRGFSEEDIAKFGFGY	
VPKRSSISIEVAEGMNITLEELVRYGIALKKGDRFVDRFEGRIVVPIKND	200
SGHIVAFGGRALGNEEPKYLNSPETRYFSKKKTFLFLDEAKKVAKEVGFF	
VITEGYFDALAFRKDGIPTAVAVLGASLSREAILKLSAYSKNVILCFDND	300
KAGFRATLKSLEDLLDYEFNVLVATPSPYKDPDELFOKEGEGSLKMLKN	
SRSFEYFLVTAGEVFFDRNSPAGVRSYLSFLKGWVQKMRRKGYLKHENL	400
VNEVSSSLQIPENQILNFFESDRSNTMPVHETKSSKVYDEGRGLAYLFLN	
YEDLREKILELDLEVLEDKNAREFFKRVSLGEDLNKVIENFPKELKDWIF	500
ETIESIPPPKDPEKFLGDLSEKLKIRRIERRIAEIDDMIKKASNDEERRL	
LLSMKVDLLRKIKRR	565

FIG. 71

ATGGCTCTACACCCGGCTCACCTGGGGCAATAATCGGGCACGAGGCCGT	
TCTCGCCCTCCTTCCCCGCCTACCGCCCAGACCCTGCTCTTCTCCGGCC	100
CCGAGGGGGTGGGGCGGCGCACCGTGGCCCGCTGGTACGCCTGGGGGCTC	
AACCGCGGCTTCCCCCGCCCTCCCTGGGGGAGCACCGGACGTCCTCGA	200
GGTGGGGCCCAAGGCCCGGGACCTCCGGGGCCGGGCCGAGGTGCGGCTGG	
AGGAGGTGGCGCCCTCTTGGAGTGGTGCTCCAGCCACCCCGGGAGCGG	300
GTGAAGGTGGCCATCCTGGACTCGGCCACCTCCTCACCGAGGCCGCGC	
CAACGCCCTCCTCAAGCTCCTGGAGGAGCCCCCTTCTACGCCCGCATCG	400
TCCTCATCGCCCCAAGCCGCGCCACCTCCTCCCCACCCCTGGCCTCCCGG	
GCCACGGAGGTGGCATTGCCCCCGTGCCCGAGGAGGCCCTGCGCGCCCT	500
CACCCAGGACCCGGAGCTCCTCCGCTACGCCGCGGGGGCCCCGGGCCG	
TCCTTAGGGCCCTCCAGGACCCGGAGGGGTACCGGGCCCCGCATGGCCAGG	600
GCGCAAAGGGTCTTGAAAGCCCCGCCCTGGAGCGCCTCGCTTTGCTTCG	
GGAGCTTTTGGCCGAGGAGGAGGGGGTCCACGCCCTCCACGCCGTCTTAA	700
AGCGCCCGGAGCACCTCCTTGCCCTGGAGCGGGCGGGGAGGCCCTGGAG	
GGGTACGTGAGCCCCGAGCTGGTCTCGCCCGGCTGGCCTTAGACTTAGA	800
GACA	

FIG. 72

MALHPAHPGAIIIGHEAVLALLPRLTAQTLLFSGPEGVGRRTVARWYAWGL	
NRGFPPPSLGEHPDVLEVGPKARDLRGRAEVRLEEVAPLLEWCSSHPRER	100
VKVAILDSAHLLEAAANALLKLLLEPPSYARIVLIAPSRATLLPTLASR	
ATEVAFAPVPEEALRALTDPELLRYAAGAPGRLLRALQDPEGYRARMAR	200
AQRVLKAPPLERLALLRELLAEEEGVHALHAVLKRPEHLLALERAREALE	
GYVSPELVLARLALDLET	268

FIG. 73

ATGCTGGACCTGAGGGAGGTGGGGGAGGCGGAGTGGAAGGCCCTAAAGCC	
CCTTTTGGAAAGCGTGCCCGAGGGCGTCCCCGTCTCCTCCTGGACCCTA	100
AGCCAAGCCCCCTCCCGGGCGGCCTTCTACCGGAACCGGGAAAGGCGGGAC	
TTCCCCACCCCCAAGGGGAAGGACCTGGTGCGGCACCTGGAAAACCGGGC	200
CAAGCGCCTGGGGCTCAGGCTCCCGGGCGGGGTGGCCAGTACCTGGCCT	
CCCTGGAGGGGGACCTCGAGGCCCTGGAGCGGGAGCTGGAGAAGCTTGCC	300
CTCCTCTCCCCACCCCTCACCCCTGGAGAAGGTGGAGAAGGTGGTGGCCCT	
GAGGCCCCCCTCACGGGCTTTGACCTGGTGCGCTCCGTCCTGGAGAAGG	400
ACCCAAGGAGGCCCTCCTGCGCCTAGGCGGCCTCAAGGAGGAGGGGGAG	
GAGCCCCTCAGGCTCCTCGGGGCCCTCTCCTGGCAGTTCGCCCTCCTCGC	500
CCGGGCCTTCTTCTCCTCCGGGAAAACCCAGGCCCAAGGAGGAGGACC	
TCGCCCGCCTCGAGGCCACCCCTACGCCGCCCGCCGCGCCCTGGAGGCG	600
GCGAAGCGCCTCACGGAAGAGGCCCTCAAGGAGGCCCTGGACGCCCTCAT	
GGAGGCGGAAAAGAGGGCCAAGGGGGGAAAGACCCGTGGCTCGCCCTGG	700
AGGCGGCGGTCTCCGCCTCGCCCGTTGA	

**FIG. 74**

MVIAFTGDPFLAREALLEEARLRGLSRFTEPTPEALAQALAPGLFGGGGA	
MLDLREVGEAEWKALKPLLESVPEGVPVLLLDPKPSPSRAAFYRNRERRD	100
FPTPKGKDLVRHLENRAKRLGLRLPGGVAQYLASLEGDLEALERELEKLA	
LLSPPLTLEKVEKVVALRPPLTGFDLVRVLEKDPKEALLRLGGLKEEGE	200
EPLRLLGALSWQFALLARAFFLLRENPRPKEEDLARLEAHPYAARRALEA	
AKRLTEEALKEALDALMEAEKRAKGGKDPWLALAAVLRLAR	292

**FIG. 75**

ATGGCTCGAGGCCTGAACCGCGTTTTCTCATCGGCGCCCTCGCCACCCG	
GCCGGACATGCGCTACACCCCGGCGGGGCTCGCCATTTTGGACCTGACCC	100
TCGCCGGTCAGGACCTGCTTCTTTCCGATAACGGGGGGGAACCGGAGGTG	
TCCTGGTACCACCGGTGAGGCTCTTAGGCCGCCAGGCGGAGATGTGGGG	200
CGACCTCTTGGACCAAGGGCAGCTCGTCTTCGTGGAGGGCCGCCTGGAGT	
ACCGCCAGTGGGAAAGGGAGGGGGAGAAGCGGAGCGAGCTCCAGATCCGG	300
GCCGACTTCCGGACCCCTGGACGACCGGGGAAGAAGCGGGCGGAGGAC	
AGCCGGGGCCAGCCAGGCTCCGCGCCGCCCTGAACCAGGTCTTCCTCAT	400
GGGCAACCTGACCCGGGACCCGGAACCTCCGCTACACCCCCAGGGCACCG	
CGGTGGCCCGGCTGGGCCTGGCGGTGAACGAGCGCCGCCAGGGGGCGGAG	500
GAGCGCACCCACTTCGTGGAGGTTTCAGGCCTGGCGCGACCTGGCGGAGTG	
GGCCGCCGAGCTGAGGAAGGGCGACGGCCTTTTCGTGATCGGCAGGTTGG	600
TGAACGACTCCTGGACCAGCTCCAGCGGCGAGCGGCGCTTCCAGACCCGT	
GTGGAGGCCCTCAGGCTGGAGCGCCCCACCCGTGGACCTGCCCAGGCCTG	700
CCCAGGCCGGCGGAACAGGTCCCGCGAAGTCCAGACGGGTGGGGTGGACA	
TTGACGAAGGCTTGAAGACTTTCGCGCGGAGGAGGATTGCGCGTTTGA	800
GCACGAA	

FIG. 76

MARGLNRVFLIGALATRPDMRYTPAGLAILDLTLAQDLLLLSDNGGEPEV	
SWYHRVRLGRQAEMWGDLLDQGQLVFVEGRLEYRQWEREGEKRSELQIR	100
ADFLDPLDDRGGKRAEDSRGPRLRAALNQVFLMGNLTRDPELRYTPQGT	
AVARLGLAVNERRQGAERTHFVEVQAWRDLAEWAAELRKGDGLFVIGRL	200
VNDSWTSSSGERRFQTRVEALRLERPTRGPAQACPGRNRNSREVQTGGVD	
IDEGLEDFPPEEDLPF	266

FIG. 77

AATTCCGACATTTCAATTGAATCGTTTATTCCGCTTGAAAAAGAAGGCAA	
GTTGCTCGTTGATGTGAAAAGACCGGGGAGCATCGTACTGCAGGCGCGCT	100
TTTTCTCTGAAATCGTGAAAAAAGTCCGCAACAAACGGTGGAATCGAA	
ACGGAAGACAACTTTTTGACGATCATCCGCTCGGGGCACTCAGAATTCGG	200
CCTCAATGGGCTAAACGCCGACGAATATCCGCGCCTGCCGCAAATTGAAG	
AAGAAAACGTGTTTCAAATCCCGGCTGATTTATTGAAAACCGTGATTCGG	300
CAAACGGTGTTTCGCCGTTTCTACATCGGAAACGCGCCCAATCTTGACAGG	
TGTCAACTGGAAAGTTGAACATGGCGAGCTTGTCTGCACAGCGACCGACA	400
GTCATCGCTTAGCCATGCGCAAAGTGAAAATTGAGTCGGAAAATGAAGTA	
TCATACAACGTCGTCATCCCTGGAAAAAGTCTTAATGAGCTCAGCAAAT	500
TTTGATGACGGCAACCACCCGGTGGACATCGTCATGACAGCCAATCAAG	
TGCTATTTAAGGCCGAGCACCTTCTCTTCTTTTCCCGGCTGCTTGACGGC	600
AACTATCCGGAGACGGCCCGCTTGATTCCAACAGAAAGCAAAACGACCAT	
GATCGTCAATGCAAAGAGTTTCTGCAGGCAATCGACCGAGCGTCCTTGC	700
TTGCTCGAGAAGGAAGGAACAACGTTGTGAAACTGACGACGCTTCCTGGA	
GGAATGCTCGAAATTTCTTCGATTTCTCCGAGATCGGGAAAGTGACGGAG	800
CAGCTGCAAACGGAGTCTCTTGAAGGGGAAGAGTTGAACATTTCGTTTCA	
CGCGAAATATATGATGGACGCGTTGCGGGCGCTTGATGGAACAGACATTT	900
CAAATCAGCTTCACTGGGGCCATGCGGCCGTTCTGTGCGCCCGCTTCA	
ACCGATTGATGCTTCAGCTCATTTTGCCGGTGAGAACATAT	992

FIG. 78

NSDISIIIESFIPLEKEGKLLVDVKRPGSIVLQARFFSEIVKKLPQQTVEI	
ETEDNFLTIIRSGHSEFRLNGLNADEYPRLPQIEEENVFQIPADLLKTVI	100
RQTVFAVSTSETRPILTVGNWVKEHGEVLCTATDSHRLAMRKVKIIESEN	
EVSYNVVIPGKSLNELSKIILDDGNHPVDIVMTANQVLFKAEHLLFFSRL	200
LDGNYPETARLIPTESKTTMIVNAKEFLQAI DRASLLAREGRNNVVKLTT	
LPGGMLEISSISPEIGKVTEQLQTESLEGEELNISFSAKYMMDALRALDG	300
TDIQISFTGAMRPFLRLPLHTDSMLQLILPVRTY	

FIG. 79

ATGATTAAACCGCGTCATTTTGGTCGGCAGGTAAACGAGAGATCCGGAGTT	
GCGTTACACTCCAAGCGGAGTGGCTGTTGCCACGTTTACGCTCGCGGTCA	100
ACCGTCCGTTTACAAATCAGCAGGGCGAGCGGGAAACGATTTTATTCAA	
TGTGTCGTTTGGCGCCGCCAGGCGGAAAACGTCGCCAACTTTTGA AAAA	200
GGGGAGCTTGGCTGGTGTGTCGATGGCCGACTGCAAACCCGCAGCTATGAAA	
ATCAAGAAGGTCGGCGTGTGTACGTGACGGAAGTGGTGGCTGATAGCGTC	300
CAATTTCTTGAGCCGAAAGGAACGAGCGAGCAGCGAGGGGCGACAGCAGG	
CGGCTACTATGGGGATCCATTCCCATTCGGGCAAGATCAGAACCACCAAT	400
ATCCGAACGAAAAAGGGTTTGGCCGCATCGATGACGATCCTTTCGCCAAT	
GACGGCCAGCCGATCGATATTTCTGATGATGATTTGCCGTTT	492

**FIG. 80**

MINRVILVGRLTRDPELRYTPSGVAVATFTLAVNRPFTNQSYENQEGRRV	
YVTEVVADSVQFLEPKGTSEQRGATAGGYQGERETDFIQCVVWRRQAEN	100
VANFLKKGSLAGVDGRLQTRGDPFPFGQDQNHQYPNEKGFGRIDDDPFAN	
DGQPIDISDDDLPF	164

**FIG. 81**

ATGCTGGAACGCGTATGGGGAAACATTGAAAAACGGCGTTTTTCTCCCCT	
TTATTTATTATACGGCAATGAGCCGTTTTTATTAACGGAAACGTATGAGC	100
GATTGGTGAACGCAGCGCTTGGCCCCGAGGAGCGGGAGTGGAACCTGGCT	
GTGTACGACTGCGAGGAAACGCCGATCGAGGCGGCGCTTGAGGAGGCCGA	200
GACGGTGCCGTTTTTTCGGCGAGCGGCGTGTCAATTCTCATCAAGCATCCAT	
ATTTTTTTTACGTCTGAAAAAGAGAAGGAGATCGAACATGATTTGGCGAAG	300
CTGGAGGCGTACTTGAAGGCGCCGTCGCCGTTTTTCGATCGTCGTCTTTTT	
CGCGCCGTACGAGAAGCTTGATGAGCGAAAAAAATTACGAAGCTCGCCA	400
AAGAGCAAAGCGAAGTCGTCATCGCCGCCCGCTCGCCGAAGCGGAGCTG	
CGTGCTGGGTGCGGCGCCGCATCGAGAGCCAAGGGGCGCAAGCAAGCGA	500
CGAGGCGATTGATGTCCTGTTGCGGCGGGCCGGGACGCAGCTTTCGCCT	
TGGCGAATGAAATCGATAAATTGGCCCTGTTTGCCGGATCGGGCGGAACC	600
ATCGAGGCGGCGGCGGTTGAGCGGCTTGTCGCCCGCACGCCGGAAGAAAA	
CGTATTTGTGCTTGTGAGCAAGTGGCGAAGCGCGACATTCCAGCAGCGT	700
TGCAGACGTTTTTATGATCTGCTTGAAAACAATGAAGAGCCGATCAAATT	
TTGGCGTTGCTCGCCGCCCATTTCCGCTTGCTTTTCGCAAGTGAAATGGCT	800
TGCCTCCTTAGGCTACGGACAGGCGCAAATTGCTGCGGCGCTCAAGGTGC	
ACCCGTTCCGCGTCAAGCTCGCTCTTGCTCAAGCGGCCCGCTTCGCTGAC	900
GGAGAGCTTGCTGAGGCGATCAACGAGCTCGCTGACGCCGATTACGAAGT	
GAAAAGCGGGGCGGTGATCGCCGGTTGGCCGTTGAGCTGCTTCTGATGC	1000
GCTGGGGCGCCCGCCCGGCGCAAGCGGGGCGCCACGGCCGGCGG	

FIG. 82

MLERVWGNIEKRRFSPLYLLYGNEPFLLTETYERLVNAALGPEEREWNLA	
VYDCEETPIEAALAEAETVPFFGERRVILIKHPYFFTSEKEKEIEHDLAK	100
LEAYLKAPSPFSIVVFFAPYEKLDERKKITKLAKEQSEVVIAAPLAEAL	
RAWVRRRIESQGAQASDEAIDVLLRRAGTQLSALANEIDKLALFAGSGGT	200
IEAAVERLVARTPEENVFVLVEQVAKRDI PAALQTFYDLLENNEEPIKI	
LALLAAHFRLLSQVKWLASLGYGQAQIAAALKVHPFRVKLALAQAARFAD	300
GELAEAINELADADYEKSGAVDRRLAVELLLMRWGARPAQAGRHR	

FIG. 83



ATGCGATGGGAACAGCTAGCGAAACGCCAGCCGGTGGTGGCGAAAATGCT	
GCAAAGCGGCTTGAAAAAGGGCGGATTTCTCATGCGTACTTGTTTGAGG	100
GGCAGCGGGGACGGGCAAAAAAGCGGCCAGTTTGTTGTTGGCGAAACGT	
TTGTTTTGTCTGTCCCAATCGGAGTTTCCCGTGTCTAGAGTGCCGCAA	200
CTGCCGGCGCATCGACTCCGGCAACCACCTGACGTCCGGGTGATCGGCC	
CAGATGGAGGATCAATCAAAAAGGAACAAATCGAATGGCTGCAGCAAGAG	300
TTCTCGAAAACAGCGGTGAGTCGGATAAAAAAATGTACATCGTTGAGCA	
CGCCGATCAAATGACGACAAGCGCTGCCAACAGCCTTCTGAAATTTTGG	400
AAGAGCCGCATCCGGGGACGGTGGCGGTATTGCTGACTGAGCAATACCAC	
CGCCTGCTAGGGACGATCGTTTCCCGCTGTCAAGTGCTTTCGTTCCGGCC	500
GTTGCCGCGCGCAGAGCTCGCCAGGGACTTGTCGAGGAGCACGTGCCGT	
TGCCGTTGGCGCTGTTGGCTGCCCATTTGACAAACAGCTTCGAGGAAGCA	600
CTGGCGCTTGCCAAAGATAGTTGGTTTGCCGAGGCGCGAACATTAGTGCT	
ACAAATGGTATGAGATGCTGGGCAAGCCGGAGCTGCAGCTTTTGTTTTTCA	700
TCCACGACCGCTTGTTTCCGCATTTTTTGGAAAGCCATCAGCTTGACCTT	
GGACTTG	757

**FIG. 84**

MRWEQLAKRQPVVAKMLQSGLEKGRISHAYLFEGQRTGKKAASLLLAQR	
LFCLSPIGVSPCLECRNCRRIDSGNHPDVRVIGPDGGSIKKEQIEWLQQE	100
FSKTAVESDKMYIVEHADQMTTSAANSLLKFLEEPHPGTVAVLLTEQYH	
RLLGTVSRQVLSFRPLPPAELAQGLVEEHVPLPLALLAAHLTNSFEEA	200
LALAKDSWFAEARTLVLQWYEMLGKPELQLLFFIHDRLFPHFLESHQLDL	
GL	252

**FIG. 85**

GTGGCATACCAAGCGTTATATCGCGTGTTTCGGCCGCAGCGCTTTGCGGA	
CATGGTTCGGCCAAGAACACGTGACCAAGACGTTGCAAAGCGCCCTGCTTC	100
AACATAAAATATCGCACGCTTACTTATTTTCCGGCCCGCGCGGTACAGGA	
AAAACGAGCGCAGCGAAAATTTTCGCCAAGGCGGTCAACTGTGAACAGGC	200
GCCAGCGGCGGAGCCATGCAATGAGTGTCCAGCTTGCCTCGGCATTACGA	
ATGGAACGGTTCCCGATGTGCTGGAAATTGACGCTGCTTCCAACAACCGC	300
GTCGATGAAATTCGTGATATCCGTGAGAAGGTGAAATTTGCGCCAACGTC	
GGCCCGCTACAAAGTGTATATCATCGACGAGGTGCATATGCTGTCGATCG	400
GTGCGTTTAACGCGCTGTTGAAAACGTTGGAGGAGCCGCCGAAACACGTC	
ATTTTCATTTTGGCCACGACCGAGCCGCACAAAATTCGGCGACGATCAT	500
TTCCCGCTGCCAACGGTTCGATTTTCGCCGCATCCCGCTTCAGGCGATCG	
TTTCACGGCTAAAGTACGTGCAAGCGCCCAAGGTGTCGAGGCGTCAGAT	600
GAGGCATTGTCCGCCATCGCCCGTGCTGCAGACGGGGGGATGCGCGATGC	
GCTCAGCTTGCTTGATCAAGCCATTTTCGTTACGCGACGGGAAACTTCGGC	700
TCGACGACGTGCTGGCGATGACCGGGGCTGCATCATTTGCCGCCTTATCG	
AGCTTCATCGAAGCCATCCACCGCAAAGATACAGCGCGCGTTCTTCAGCA	800
CTTGGAACGATGATGGCGCAAGGGAAAGATCCGCATCGTTTGTTGGAAG	
ACTTGATTTTGTACTATCGCGATTTATTGCTGTACAAAACCGCTCCCTAT	900
GTGGAGGGAGCGATTCAAATTGCTGTGCTTGACGAAGCGTTCACTTCACT	
GTCGGAAATGATTCCGGTTTCCAATTTATACGAGGCCATCGAGTTGCTGA	1000
ACAAAAGCCAGCAAGAGATGAAGTGGACAAACCACCGCGCCTTCTGTTG	
GAAGTGGCGCTTGTGAAACTTTGCCATCCATCAGCCGCCGCCCGTCGCT	1100
GTCGGCTTCGAGTTGGAACCGTTGATAAAGCGGATTGAAACGCTGGAGG	
CGGAATTGCGGCGCCTGAAGGAACAACCGCCTGCCCCCTCCGTCGACCGCC	1200
GCGCCCGTGAAAAAACTGTCCAAACCGATGAAAACGGGGGGATATAAAGC	
CCCGGTTGGCCGCATTTACGAGCTGTTGAAACAGGCGACGCATGAAGATT	1300
TAGCTTTGGTGAAAGGATGCTGGGCGGATGTGCTCGACACGTTGAAACGG	
CAGCATAAAGTGTGCGACGCTGCCTTGCTGCAAGAGAGCGAGCCGGTTGC	1400
AGCGAGCGCCTCAGCGTTTGTATTAAATTCAAATACGAAATCCACTGCA	
AAATGGCGACCGATCCCAAGTTCGGTCAAAGAAAACGTCGAAGCGATT	1500
TTGTTTGAGCTGACAAACCGCCGCTTTGAAATGGTAGCCATTCCGGAGGG	
AGAATGGGGAAAAATAAGAGAAGAGTTCATCCGCAATAAGGACGCCATGG	1600
TGGAAAAAAGCGAAGAAGATCCGTTAATCGCCGAAGCGAAGCGGCTGTTT	
GGCGAAGAGCTGATCGAAATTAAAGAA	1677

FIG. 86

VAYQALYRVFRPQRFADMVGQEHVTKTLQSALLQHKISHAYLFSGPRGTG	
KTSAAKIFAKAVNCEQAPAAEPCNECPACLGITNGTVPDVLEIDAASNNR	100
VDEIRDIREKVKFAPTSARYKVYIIDEVHMLSIGAFNALLKTLEPPKHV	
IFILATTEPHKIPATIIISRCQRFDFRRIPLQAIVSRLKYVASAQGVEASD	200
EALSAIARAADGGMRDALSLDQAI SFSDGKLRLDDVLAMTGAASFAALS	
SFIEAIHRKDTAAVLQHLETMMAQGKDPHRLVEDLILYYRDLLEYKTAPY	300
VEGAIQIAVVDEAFTSLSEMI PVSNLYEAI ELLNKSQQEMKWTNHPRLLL	
EVALVKLCHPSAAAPSL SASELEPLIKRIETLEAELRRLKEQPPAPPSTA	400
APVKKLSKPMKTGGYKAPVGRIYELLKQATHEDLALVKGCWADVLDLTKR	
QHKVSHAALLQESEPVAASASAFVLKFYIEIHCKMATDPTSSVKENVEAI	500
LFELTNRRFEMVAIPEGEWGKIREEFIRNKDAMVEKSEEDPLIAEAKRLF	
GEELIEIKE	559

**FIG. 87**

ATGGTGACAAAAGAGCAAAAAGAGCGGTTTCTCATCCTGCTTGAGCAGCT	
GAAGATGACGTCCGACGAATGGATGCCGCATTTTCGTGAGGCAGCCATTC	100
GCAAAGTCGTGATCGATAAAGAGGAGAAAAGCTGGCATTTTTATTTTCAG	
TTCGACAACGTGCTGCCGGTTCATGTATACAAAACGTTTGCCGATCGGCT	200
GCAGACGGCGTTCCGCCATATCGCCGCGTCCGCCATACGATGGAGGTCG	
AAGCGCCGCGCGTAAGTGAAGGCGGATGTGCAGGCGTATTGGCCGCTTTGC	300
CTTGCCGAGCTGCAAGAAGGCATGTCGCCGCTTGTCGATTGGCTCAGCCG	
GCAGACGCCTGAGCTGAAAGGAAACAAGCTGCTTGTCGTTGCCCCGCCATG	400
AAGCGGAAGCGCTGGCGATCAAACGGCGGTTCCGCCAAAAAATCGCTGAT	
GTGTACGCTTCGTTTGGGTTTCCCCCCTTCAGCTTGACGTCAGCGTCGA	500
GCCGTCCAAGCAAGAAATGGAACAGTTTTTGGCGCAAAAACAGCAAGAGG	
ACGAAGAGCGAGCGCTTGCTGTACTGACCGATTTAGCGAGGGAAGAAGAA	600
AAGGCCGCGTCTGCGCCGCGTCCGGTCCGCTTGTCATCGGCTATCCGAT	
CCGCGACGAGGAGCCGGTGCGGCGGCTTGAAACGATCGTCGAAGAAGAGC	700
GGCGCGTCGTTGTGCAAGGCTATGTATTTGACGCCGAAGTGAGCGAATTA	
AAAAGCGGCCGACGCTGTTGACCATGAAAATCACAGATTACACGAATC	800
GATTTTAGTCAAAATGTTCTCGCGCGACAAAGAGGACGCCGAGCTTATGA	
GCGGCGTCAAAAAGGCATGTGGGTGAAAGTGCGCGGCAGCGTGCAAAAC	900
GATACGTTTCGTCCGTGATTTGGTCATCATCGCCAACGATTTGAACGAAAT	
CGCCGCAACGAACGGCAAGATACGGCGCCGGAAGGGGAAAAGAGGGTCG	1000
AGTCCATTTGCATACCCCGATGAGCCAAATGGACGCGGTACCTCGGTG	
ACAAACTCATTGAGCAAGCGAAAAAATGGGGGCATCCGGCGATCGCCGT	1100
CACCGACCATGCCGTTGTTTCAGTCGTTTCCGGAGGCCTACAGCGCGGCGA	
AAAAACACGGCATGAAGGTCATTTACGGCCTTGAGGCGAACATCGTCGAC	1200
GATGGCGTGCCGATCGCCTACAATGAGACGCACCGCCGTCTTTCGGAGGA	
AACGTACGTCGTCTTTGACGTCGAGACGACGGGCCTGTCGGCTGTGTACA	1300
ATACGATCATTGAGCTGGCGGCGGTGAAAGTGAAAGACGGCGAGATCATC	
GACCGATTTCATGTCGTTTGCCAACCCTGGACATCCGTTGTCGGTGACAAC	1400
GATGGAGCTGACTGGGATCACCGATGAGATGGTGAAAGACGCCCCGAAGC	
CGGACGAGGTGCTAGCCCGTTTTGTTGACTGGGCGGCGATGCGACGCTT	1500
GTTGCCACAAACGCCAGCTTTGACATCGGTTTTTTAAACGCGGGCCTCGC	
TCGCATGGGGCGCGGCAAAATCGCGAATCCAGTCATCGATACGCTCGAGC	1600
TGGCCCGTTTTTTTATACCCGGATTTGAAAACCATCGGCTCAATACATTG	
TGCAAAAATTTGACATTGAATTGACGCAGCATCACCGCGCCATCTACGA	1700
CGCGGAGGCGACCGGGCATTGCTTATGCGGCTGTTGAAGGAAGCGGAAG	
AGCGCGGCATACTGTTTCATGACGAATTAAACAGCCGCACGCACAGCGAA	1800
GCGTCCATATCGGCTTGCGCGCCCGTTCATGTGACGCTGTTGGCGCAAAA	
CGAGACTGGATTGAAAAATTTGTTCAAGCTTGTTGTCATTGTGCGACATTC	1900
AATATTTTACCCTGTGCCGCGCATCCCGCGCTCCGTGCTCGTCAAGCAC	
CGCGACGGCCTGCTTGTCGGCTCGGGCTGCGACAAAGGAGAGCTGTTTGA	2000
CAACTTGATCCAAAAGGCGCCGGAAGAAGTCGAAGACATCGCCCGTTTTT	
ACGATTTTCTTGAAGTGCATCCGCCGACGTGTACAAGCCGCTCATCGAG	2100
ATGGATTATGTGAAAGACGAAGAGATGATCAAAAACATCATCCGCAGCAT	
CGTCGCCCTTGGTGAGAAGCTTGACATCCCGGTTGTCGCCACTGGCAACG	2200

FIG. 88A

TCCATTACTTGAACCCAGAAGATAAAATTTACCGGAAAATCTTAATCCAT  
TCGCAAGGCGGGGCGAATCCGCTCAACCGCCATGAACTGCCGGATGTATA 2300  
TTTCCGTACGACGAATGAAATGCTTGACTGCTTCTCGTTTTTAGGGCCGG  
AAAAAGCGAAGGAAATCGTCGTTGACAACACGCAAAAAATCGCTTCGTTA 2400  
ATCGGCGATGTCAAGCCGATCAAAGATGAGCTGTATACGCCGCGCATTGA  
AGGGGCGGACGAGGAAATCAGGGAAATGAGCTACCGGCGGGCGAAGGAAA 2500  
TTTACGGCGACCCGTTGCCGAAACTTGTTGAAGAGCGGCTTGAGAAGGAG  
CTAAAAGCATCATCGGCCATGGCTTTGCCGTCATTTATTTGATCTCGCA 2600  
CAAGCTTGTGAAAAATCGCTCGATGACGGCTACCTTGTCGGGTCGCGCG  
GATCGGTCGGCTCGTCGTTTGTCGCGACGATGACGGAAATCACCGAGGTC 2700  
AATCCGCTGCCGCCGATTACGTTTGCCCGAACTGCAAGCATTTCGGAGTT  
CTTTAACGACGGTTCAGTCGGCTCAGGGTTTGATTTGCCGGATAAAAACT 2800  
GCCCCGCGATGTGGGACGAAATACAAGAAAGACGGGCACGACATCCCGTTT  
GAGACGTTTCTCGGCTTTAAAGGCGACAAAGTGCCGGATATCGACTTGAA 2900  
CTTTTCCGGCGAATACCAGCCGCGCGCCCACTATACGAAAGTGCTGT  
TTGGCGAAGACAACGTCTACCGCGCCGGGACGATTGGCACGGTCGCTGAC 3000  
AAAACGGCGTACGGATTTGTCAAAGCGTATGCGAGCGACCATAACTTAGA  
GCTGCGCGGCGCGGAAATCGACGGCTCGCGGCTGGCTGCACCGGGGTGAA 3100  
GCGGACGACCGGGCAGCATCCGGGCGGCATCATCGTCGTCCCGGATTATA  
TGGAATTTACGATTTTACGCCGATTCAATATCCGGCCGATGACACGTCC 3200  
TCTGAATGGCGGACGACCCATTTGACTTCCATTTCGATCCACGACAATTT  
GTTGAAGCTCGATATTCTCGGGCAGCAGATCCGACGGTCATTTCGCATGC 3300  
TGCAAGATTTAAGCGGCATCGATCCGAAAACGATCCCGACCGACGACCCG  
GATGTGATGGGCATTTTCAGCAGCACCGAGCCGCTTGCGGTTACGCCGGA 3400  
GCAAATCATGTGCAATGTGCGCACGATCGGCATTCCGGAGTTTGGCACGC  
GCTTCGTTCCGGCAAATGTTGGAAGAGACAAGGCCAAAAACGTTTTCCGAA 3500  
CTCGTGCAAATTTCCGGCTTGTCGCACGGCACCGATGTGTGGCTCGGCAA  
CGCGCAAGAGCTCATTCAAAACGGCACGTGTACGTTATCGGAAGTCATCG 3600  
GCTGCCGCGACGACATTATGGTCTATTTGATTTACCGCGGGCTCGAGCCG  
TCGCTCGCTTTTAAAATCATGGAATCCGTGCGCAAAGGAAAAGGCTTAAC 3700  
GCCGGAGTTTGAAGCAGAAATGCGCAAACATGACGTGCCGGAGTGGTACA  
TCGATTCATGCAAAAAATCAAGTACATGTTCCCGAAAGCGCACGCCGCC 3800  
GCCTACGTGTTAATGGCGGTGCGCATCGCCTACTTTAAGGTGCACCATCC  
GCTTTTGTATTACGCGTCGTACTTTACGGTGCGGGCGGAGGACTTTGACC 3900  
TTGACGCCATGATCAAAGGATCACCCGCCATTTCGCAAGCGGATTGAGGAA  
ATCAACGCCAAAGGCATTCAGGCGACGGCGAAAGAAAAAGCTTGCTCAC 4000  
GGTTCTTGAGGTGGCCTTAGAGATGTGCGAGCGCGGCTTTTCCTTTAAAA  
ATATCGATTTGTACCGCTCGCAGGCGACGGAATTCGTCATTGACGGCAAT 4100  
TCTCTCATTCCGCCGTTCAACGCCATTCCGGGGGCTTGGGACGAACGTGGC  
GCAGGCGATCGTGCGCGCCCGCGAGGAAGGCGAGTTTTTGTGCAAGGAGG 4200  
ATTTGCAACAGCGCGGCAAATTGTGCAAAACGCTGCTCGAGTATCTAGAA  
AGCCGCGGCTGCCTTGACTCGCTTCCAGACCATAACCAGCTGTCGCTGTT 4300  
T

FIG. 88B

MVTKEQKERFLILLEQLKMTSDEWMPHFREAAIRKVVIDKEEKSWHFYFQ	
FDNVLPVHVYKTFADRLQTAFRHIAAVRHTMEVEAPRVTEADVQAYWPLC	100
LAEHQEGMSPLVDWLSRQTPPELKGNKLLVVARHEAEALAIKRRFAKKIAD	
VYASFGFPPLQLDVSVEPSKQEMEQLAQKQQEEDERALAVLTDLAREEE	200
KAASAPPSGPLVIGYPIRDEEPVRRLETIVEEERRVVVQGYVFDAEVSEL	
KSGRTLTTMKITDYTNLSILVKMFSRDKEDAELMSGVKKGMWVKVGRSVQN	300
DTFVRDLVIIANDLNEIAANERQDTAPEGEKRVELHLHTPMSQMDAVTSV	
TKLIEQAKKWGHFAIAVTDHAVVQSFPEAYSAAKKHGMKVIYGLEANIVD	400
DGVPIAYNETHRRLSEETYVVFVETTGLSAVYNTIIELA AVKVKDGEII	
DRFMSFANPGHPLSVTTMELTGITDEMVKDAPKPDEV LARFVDWAGDATL	500
VAHNASFDIGFLNAGLARMGRGKIANPVIDTLELARFLYPDLKNHRLNTL	
CCKFDIELTQHHRAIYDAEATGHLLMRLKKEAEERGILFHDELNSRTHSE	600
ASYRLARPFHVTLQAQNETGLKNLFKLVSLSHIQYFHRVPRI PRSVLVKH	
RDGLLVGSGCDKGELFDNLIQKAPEEVEDIARFYDFLEVHPPDVYKPLIE	700
MDYVKDEEMIKNIIRSIVALGEKLDIPVVATGNVHYLNPEDKIYRKILIH	
SQGGANPLNRHELDPVYFRTTNEMLD CFSFLGPEKAKEIVDNTQKIASL	800
IGDVKPIKDELYTPRIEGADEEIREMSYRRAKEIYGDPLPKLVEERLEKE	
LKSIIGHGFAVIYLI SHKLVKKSLLDGYLVGSRGSSSVFVATMTEITEV	900
NPLPPHYVCPNCKHSEFFNDGSGVSGFDLPDKNCPRCGTYKKDGHDI PF	
ETFLGFKGDKVPDIDLNFSGEYQ PRAHNYTKVLFGEDNVYRAGTIGTVAD	1000
KTAYGFVKAYASDHNLELRGAEIDLAAGCTGVKRTTGQHPGGIIVVPDYM	
EIYDFTPIQYPADDTSSSEWRTHFDFHSIHDNLLKLDILGHDDPTVIRML	1100
QDLSGIDPKTIPTDDPDVMGIFSSTEPLGVTPEQIMCNVGTIGIPEFGTR	
FVRQMLEETRPKTFSELVQISGLSHGTDVWLGN AQELIQNGTCTLSEVIG	1200
CRDDIMVYLIYRGLEPSLAFKIMESVRKGKGLTPEFEAEMRKHDVPEWYI	
DSCKKIKYMFPAHAHAAAYVLM AVRIAYFKVHHPLLYASYFTVRAEDFDL	1300
DAMIKGSPAIRKRIEEINAKGIQATAKEKSLLT VLEVALEM CERGFSFKN	
IDLYRSQATEFVIDGNSLIPPFNAIPGLGTNVAQAIVRAREEGEFLSKED	1400
LQQRGKLSKTLLEYLESRGCLDSL PDHNQLSLF	

FIG. 89